

THE PHONETICS AND PHONOLOGY OF
THE KARANGA DIALECT OF SHONA AS
SPOKEN IN THE MIDLANDS REGION
OF RHODESIA

A thesis submitted for the degree
of Doctor of Philosophy in the
University of London by

ALEC JOSHUA COSMO PONGWENI

of the School of Oriental and African Studies.

Supervisor: Professor J. Carnochan.

October, 1977.



ProQuest Number: 10731305

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10731305

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

This thesis is in two parts. Part I is a phonetic study of the sounds of the Karanga dialect of Shona, prefaced by a review of some of the major works on Shona phonetics. The review, written from the point of view of current scientific theory, isolates those areas of the subject in which the present author hopes to make a contribution.

In order to bring some order to the analysis and classification of the sounds of the dialect, a brief outline of the phoneme theory as set out by the post-Bloomfieldians is given in the first chapter. The review is thus set against the background of that theory and it is argued that the work of Doke (1931) and Fortune (1955, 1968) would have benefitted if it had been conducted within the framework of an explicit theory such as that propounded by the American structuralist phonemicists.

Chapters 3, 4 and 5 are a phonetic study of the sounds of Karanga: the distinctive segments of the dialect are grouped into classes using articulatory criteria. Then each class is described with a view to pin-pointing some of the outstanding acoustic features¹ which characterise it and such features are compared with those reported for the sounds of other languages in articles and books by other phoneticians.

The outstanding phonetic features described in Part I are used as the basis for an inventory of distinctive features used in Part II. In the latter part, in

1. All the instrumental material is contained in Vol.II of the thesis.

particular in Chapter 7, the distinction between phonology and grammar has become very narrow indeed. That Chapter (Chapter 6 is on segmental phonological processes) is on the tonal processes caused by the affixation of certain grammatical prefixes on to nominal stems, e.g. what have been called the prefixes of "association" and "similarity", /nà-/, "with" or "in the company of..", and /sà-/, "like" or "resembling", respectively.

The nouns of Shona fall into a number of classes each of which is made up of items with a variety of tone patterns. The study reported in that chapter is interesting in that the rules which one can posit to account for these tonal processes are, to a large extent, similar, in spite of the fact that the prefixes perform diverse grammatical functions.

CONTENTS

	<u>Page</u>
ABSTRACT	2
CONTENTS	4
Preface and Acknowledgements	12
<u>CHAPTER 1</u>	
1.0 A Review of the Literature on Shona Phonetics	18
1.1 Introduction.	18
1.2 On the nature of scientific theories	19
1.3 The American Structuralists and Phonemic Analysis	30
1.4 <u>A Comparative Study in Shona Phonetics</u>	42
1.4.1 <u>An Analytical Grammar of Shona</u>	58
1.4.2 <u>Shona Grammatical Constructions</u>	68
1.4.2.1 Fortune's Phonemic analysis of the stops	77
1.4.2.2 An Alternative Phonemic Analysis of the stops	78
1.4.2.3 Fortune's analysis of Suprasegmental phenomena	84
<u>CHAPTER 2</u>	
2.0 Phonetic Analysis and Classification: a theoretical framework	92
2.1 On the "Speech Chain"	92
2.2 The Physiological Phase	93
2.2.1 The Major Initiation Types	94
2.3 The Articulatory Stage	97
2.3.1 The Condition of Cavity	99
2.3.2 The degree of stricture	101
2.3.3 Articulatory Potential	102
2.4 Modes of egress	103
2.5 Minor Articulatory Features	104
2.6 The Auditory Stage	107

	Page
2.6.1 Major auditory phenomena	107
2.6.2 Minor auditory phenomena	108
2.7.0 Continuant and Non-Continuant	109
2.7.1 Continuants	109
2.7.2 Non-Continuants	110
2.8.0 A Note on Tone	112
2.9.0 The Acoustic Phase	114
2.9.1 Some Acoustic Concepts	115
2.9.2 Conclusion	118
<u>CHAPTER 3</u>	
3.0 Description and classification of the Vowels of Karanga	119
3.1 The Vowels of Karanga in Isolation	124
3.2 The hV complexes	128
3.2.1 The vowel sequences	137
3.2.2 Description of the vowel sequences	139
3.2.3 Interpretation of the Spectrograms of the vowel sequences	143
3.2.4 The grammatical function of the vowels in the sequences	156
3.2.5 Sequences of Vowels within a lexical item	163
3.2.6 The distribution of glides among vowel sequences	166
3.2.7 (a) Examples that allow an optional [j] glide	168
(b) Examples that allow an optional [w] glide	168
3.2.8 Examples allowing the insertion of neither glide	169
3.3 Conclusion	170
3.3.1 The phonological argument	170
3.3.2 The syntactic argument	171

CHAPTER 4

The Semi-Vowels	175
4.0 Introduction	175
4.1 The labio-velar semi-vowels /w/ and /w̥/.	176
4.2 The palatal semi-vowel /j/.	177
4.3.0 Acoustic Studies of semi-vowels.	180
4.3.1 The number of Formants Needed	182
4.3.2 Duration of Steady-state Onsets	184
4.3.3 The Starting Frequencies of the Transitions	184
4.3.4 Transition Duration	184
4.4.0 Acoustic Features of Karanga Semi-vowels	186
4.4.1 The labio-dental frictionless continuant, /ʋ/.	192

CHAPTER 5

The Consonantal Phonemes of Karanga	198
5.0 Introduction	198
5.1 Choosing Examples for Palatography	200
5.2 The Palatogram Figure	202
5.3 The Plosive Phonemes, Introduction	203
5.3.1 The Search for Acoustic Cues	206
5.3.2 The Plosive Phonemes	211
5.3.3 The bilabial explosives, /p, p̥/.	213
5.3.4 The Alveolar explosives, /t, t̥/.	217
5.3.5 The Velar explosives, /k, k̥/.	222
5.3.6 The bilabial implōsive, /b̥/.	226
5.3.7 The alveolar implōsive, /d̥/.	229
5.3.8 The alveolar click /ʄ/.	232
5.3.9 The Phonemic Interpretation of the Stop Consonants	235
5.3.10.0 The Fricative Phonemes	240
5.3.10.1 Introduction	240

	<u>Page</u>
5.3.10.2 The Search for Acoustic Cues	243
5.3.10.3 The labio-dental fricatives, /f, v/.	247
5.3.10.4 The dental fricatives, /s, z/.	249
5.3.10.5 The labialised alveolar fricatives, /ʃ, ʒ/.	252
5.3.10.6 The alveolar lateral fricatives, /ɬ, ɮ/.	255
5.3.10.7 The palato-alveolar fricatives, /ʃ, ʒ/.	257
5.3.10.8 The velar fricatives, /x, ɣ/.	260
5.3.10.9 The glottal fricative, /h/.	262
5.3.10.10 Conclusion	264
5.3.10.11 The Affricates of Karanga.	266
5.3.10.12 The labio-dental affricates, /pf, bʏ/.	268
5.3.10.13 The denti-alveolar affricates, /ts, dʒ/.	271
5.3.10.14 The labialised alveolar affricates, /tʃ, dʒ/.	274
5.3.10.15 The palato-alveolar affricates, /tʃ, dʒ/.	277
5.3.10.16 The labio-velar "affricates" /px, bɣ/.	279
5.3.11.0 The Nasal Phonemes	282
5.3.11.1 Acoustic Studies of Nasals	285
5.3.11.2 The bilabial nasals, /m, m̥/.	286
5.3.11.3 The alveolar nasals, /n, n̥/.	290
5.3.11.4 The palatal nasals, /ɲ, ɲ̥/.	292
5.3.11.5 The velar nasals, /ŋ, ŋ̥/.	294
5.3.12.0 The Naso-oral Phonemes.	298
5.3.12.1 The Labial naso-oral phonemes	302
5.3.12.2 The alveolar naso-oral phonemes	308
5.3.12.3 The breathy-voiced palato-alveolar naso-oral phoneme /ɲɔ̥/	317
5.3.12.4 The velar naso-oral phonemes	319
5.3.12.5 The labial-velar naso-oral phoneme, /mɔ̥ɣ/.	321
5.3.13.0 The rolled alveolar phoneme /r/.	323

5.3.13.1	The Labialised phonemes of Karanga	325
5.3.13.2	The labialised bilabial nasal, /m ^w /.	327
5.3.13.3	The labialised velar phonemes	328
5.3.13.4	The labialised velar plosives /k ^w , g ^w /.	328
5.3.13.5	The labialised velar nasals /ŋ ^w , ɲ ^w /.	330
5.3.13.6	The labialised velar naso-oral phoneme /ŋg ^w /.	331
5.3.13.7	The labialised velar fricatives /x ^w , ɣ ^w /.	332
<u>CHAPTER 6</u>		
6.0	The Segmental Phonology of Karanga	335
6.1	Introduction	335
6.1.1	Phonology and other levels of analysis	335
6.1.2	Dependence of phonology on syntax	336
6.1.3	Dependence of phonology on phonetics	337
6.1.4	Levels of sound representation	338
6.1.5	Linguistically significant generalizations	339
6.1.6	The structure of a transformational generative grammar	341
6.1.7	The function of the phonological component	342
6.1.8	A set of distinctive features for the sounds of Karanga	342
6.2.0	The Distinctive features of Chomsky and Halle	345
6.2.1	Problems relating to the consonantal phonemes of Karanga	349
6.2.2	Some phonological processes in Karanga	354
6.2.3	The lexical representation of Karanga words	355
6.3.0	Word derivation processes	357
6.3.1	The derivation of class 9/10 nouns from verbs	359
6.3.2	The derivation of class 5 nouns from /p, t, k/- commencing verb stems	369
6.3.3	The derivation of plural forms	371

	<u>Page</u>
6.3.3.1 Prefixes in primary and secondary function	372
6.3.3.2 Plurals formed by vowel alternation	373
6.3.4.0 Plurals formed by changes in features of consonant of the Prefix	374
6.3.5 Plurals formed by the addition of a prefix	375
<u>CHAPTER 7</u>	
7.0 Some tonal processes of Karanga	385
7.1.0 Introduction	385
7.1.1 A tone language	385
7.1.2 Relative [+HIGH] and Relative [-HIGH]	388
7.1.3 Downstep and Downdrift	392
7.2.0 The consonantal margin and syllable pitch	395
7.2.1 Tonal Assimilation and Dissimilation	396
7.2.2 Downdrift in Karanga	397
7.3.0 The Adverbial prefixed of "similarity" and "association" and noun tone	398
7.3.1 The nouns of classes 1 and 2.	399
7.3.2 The nouns of class 1a.	402
7.3.3 The nouns of class 2a.	404
7.3.4 The nouns of classes 3 and 4.	407
7.3.5 The nouns of classes 5, 9/10.	409
7.4.0 The adverbial prefix of instrumentality	414
7.4.1. The stabilization construction	416
7.4.1.1 Stabilization constructions with a class prefix	417
7.4.1.2 Stabilization constructions indicated by /i'-./	424
7.4.1.3 Stabilization constructions indicated by /ndi'-./	428
7.5.0 The treatment of the object infix in Karanga.	433
7.5.1 The infinitive inflection and the object infix	434
7.5.1.1 [-HIGH] stems	434

	<u>Page</u>
Figure 4 Formant pattern of [-oa-]	151
Figure 5 Formant pattern of [-ou-]	152
Figure 6 Formants of /l,r,w,j/.	183
Figure 7 Formant transitions of [-aja-]	188
Figure 8 Formants of [-ewe-]	190
Figure 9 Formants [-ir-]	195
Figure 10 Formants [-eʋe-]	196
Figure 11: Palatogram Figure	202

APPENDICES

Appendix I The tones of derived nominals	446
Appendix II Derivation of the past tense	449
Appendix III Derivation of the future tense	452

<u>BIBLIOGRAPHY</u>	454
---------------------	-----

VOL. II

- Inside back page: 1. Sonagram grid
2. Palatogram grids (a) and (b).

PREFACE AND ACKNOWLEDGEMENTS

Rhodesia is divided, for administrative purposes, into two main regions, Mashonaland and Matabeleland. The main vernacular of the latter region is Ndebele. The inhabitants of this part of the country are historically related to the Zulus of South Africa, having come to the country in the 19th Century under the leadership of Mzilikazi and Lobengula, runaway chiefs previously under the warrior King Chaka of the Zulus.

Compared to Shona, Ndebele is a fairly homogeneous language in that there is no radical dialectal variation from area to area. Shona, on the other hand, is an amalgam of dialects and sub-dialects. It is a recent creation of the government and the Church designed to facilitate administration and educational planning. Following Doke's Report on the Unification of the Shona Dialects presented to the Legislative Assembly of Southern Rhodesia in 1931, and subsequent language committees, a standard language called "Shona" based mainly on Zezuru dialect features, but also on Karanga, Manyika and Korekore, has been developed and taught in schools. The University of Rhodesia's Department of African Languages does extensive research on all the dialects: two M.Phil. theses written in the last eight years are on the Korekore Dialect (Dembetembe, 1969) and the Ndau Dialect (Mkanganwi, 1973), and Dembetembe's (1976) Ph.D. thesis called The syntax of sentential complements in Shona is based mainly on the Zezuru dialect.

The neat picture of a homogeneous standard language

presented in school textbooks has no resemblance to the speech of people in different areas of the country. Children learn Zezuru-based standard Shona at school but speak to their parents and friends in the local dialect - see N.C. Dembetembe (1974)¹.

The language picture is further complicated by the fact that each dialect comprises a number of sub-dialects. Thus, the present author speaks Karanga, but specifically the Mhari sub-dialect of Karanga. There are certain phonetic features in his speech which other Karanga speakers from different areas find peculiar. The main unifying features among the dialects of Shona are

- (a) their use of lexical tone, i.e. they are tone "languages" - see chapter 6.

- (b) they have a two-tone system, [+HIGH] and [-HIGH].

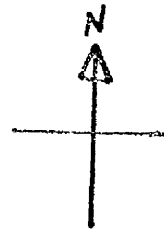
Since these relative tones are capable of bringing about semantic differences between forms otherwise composed of the same segmental units, they have phonemic status and are sometimes symbolised as /H/, /L/.

- (c) they share a considerable percentage of their vocabularies, making mutual intelligibility maximal. Their geographic proximity, enhanced by population mobility resulting from industrialization and urbanization, tend to increase rather than diminish inter-dialectal communication.

The accompanying map (page 14) is adapted from Hannan's (1974) Standard Shona Dictionary. It shows the locations of the various dialects in the Mashonaland area,

1. Dembetembe, N.C. (1974) "Three Korekore Dialects" African Language Studies, XV.

DISTRIBUTION OF THE DIALECTS OF SHONA.



(After C.M.Doke, 1931, and M.Hannan, 1974)

with the relevant sub-dialects indicated only for the Karanga dialect. This thesis is essentially an analysis of the author's speech. As a product of the post-industrial African enlightenment, he does not claim to have had no contact with people from other dialects. Many of his school-mates were from the Zezuru, Manyika and Ndau areas and he is now married to a Manyika. He has therefore listened to Zezuru, Ndau and Manyika parlance. But the examples used in this thesis are Mhari Karanga.

While the phonetic study is as comprehensive as space and time permitted, the phonological one is not at all exhaustive of the processes found in the dialect. The latter study is merely illustrative of the lines that a more extended treatment of Karanga phonology might follow.

The author considers the contribution of this study to Shona linguistics to be of two kinds: theoretical and practical. The phoneme theory has been applied to the analysis and classification of the sounds of the Karanga dialect; and in the phonology section an attempt has been made to formalise some of the phonological processes in Karanga within the generative phonology model. In the event, it was found that the set of distinctive features proposed by Chomsky and Halle (1968), though meant to apply to all languages, are quite inadequate to handle the contrasts in the Karanga sound system.

Secondly, the phonetics laboratory facilities at SOAS have provided ample opportunity for the author to isolate some of the informative acoustic cues for the perception of the sounds of various classes in Karanga.

In particular the locus theory has provided some fairly constant acoustic correlates to the physiological, i.e. articulatory, parameter of place of articulation - although it was found less helpful in distinguishing between adjacent places of articulation, e.g. bilabial from labiodental.

Finally, I would like to express my deep felt gratitude to Professor Jack Carnochan, who has, by his example taught me to cultivate a healthy eclecticism in research, and encouraged me throughout. Mr. A.W. Stone, the Phonetics Laboratory's Chief Technician, worked extremely hard, especially on the Palatograms and Spectrograms, many of which were rejected before we selected those used in the thesis. I am very grateful to him. Dr. H.J. Carter discussed a number of issues relating to tonal phonology with me and provided much valuable literature on the subject. My thanks are also due to Dr. J.C. Wells (U.C.L) Dr. D.C. Bennett and Dr. R.J. Hayward (S.O.A.S) for reading parts of this thesis and making many helpful suggestions. While acknowledging my debt to all these people, I accept full responsibility for any shortcomings in this thesis.

I am also grateful to the Foreign and Commonwealth Office for nominating me for a Commonwealth Scholarship which enabled me to study Phonetics at University College, London, (1972-74) and Linguistics at Leeds (1974-75). The M.O.D. awarded me a grant for doing research for this thesis at SOAS. To them and all my British Council Programme Organisers, my thanks are due.

Clapham North.

A.P.

October , 1977.

PART I

A PHONETIC STUDY OF THE
SOUNDS OF THE KARANGA DIALECT

CHAPTER 1

1.0 A REVIEW OF THE LITERATURE ON SHONA PHONETICS

1.1 INTRODUCTION

C.M. Doke's (1931)¹ A Comparative Study in Shona Phonetics, the object of this part of the review, "covers the more technical and scientific phonetic findings of my year's work." (preface). His equipment in the field comprised "field instruments, precluding many of the more delicate experiments which can today (1929) be made in a phonetic laboratory" (preface). Specifically he used palatography, which "was limited by expense and time", factors which "prevented more than one subject from each dialect being used" (preface). He also made use of the Kymograph, and the camera.

In order to assess Doke's achievement, I think we must place him in his intellectual milieu. What were the current attitudes towards scientific work on language? And what was their attitude towards science? How did Doke differ from other linguists, and why?

This chapter will be concerned with:-

- (i) the brief exposition of the nature of scientific theory;
 - (ii) the assessment of American Structuralist phonemics.
- The American theory of the phoneme will be assessed in terms of what seem to be the most valuable aspects of (i)

1. Doke, C.M. (1931) A Comparative Study in Shona Phonetics, University of the Witwatersand Press.

for linguistic work.

(iii) The last part of this section will be something of a practical criticism of A Comparative Study in Shona Phonetics against the background of what has been said in the preceding sections.

It is the view of the present author that every researcher must be aware of work that has been done by others in his field of specialization. By being constantly aware of the work of one's predecessors as well as that of contemporaries, one is indirectly engaging in debate with them. This view is expressed by Popper (1959, pp.16-17):²

"If we ignore what other people are thinking, or have thought in the past, then rational discussion must come to an end, though each of us may go on happily talking to himself. Some philosophers have made a virtue of talking to themselves; perhaps because they felt that there was nobody else worth talking to. I fear that the practice of philosophising on this somewhat exalted plane may be a symptom of the decline of rational discussion."

1.2 On the nature of scientific theories

There are two main branches of scientific enquiry: the empirical and the non-empirical. The empirical sciences explore, describe and attempt to explain and to predict occurrences in the world we live in. Their statements are checked against the facts of our experience and are

2. Popper, K.R. (1959) The Logic of Scientific Discovery Hutchinson of London, 8th Impression Oct. 1975.

acceptable only if they are supported by empirical evidence, which is obtained by systematic observation, experimentation, interviews and (some would add) introspection.

The perennial task of science is the search for natural laws. An important distinction has to be made between natural law on the one hand, and the laws of society on the other. Natural law describes what happens in the universe. Man cannot add to or refine that law. The laws of society, however, are prescriptive, and as such can either be obeyed or broken.

The method which scientists have traditionally adopted in their search for natural laws has the following stages. First, the scientist experiments and observes. The aim of his experiments is to make carefully controlled and meticulously measured observations. His area of research is that twilight zone between our knowledge, so-called, and our ignorance. He systematically records his findings, and may even publish them. He and other workers in the field gather a lot of data. When general patterns begin to emerge, the scientist then formulates a hypothesis concerning his data. The next step is for him to look for evidence which supports the hypothesis. This goes on for a considerable time, until the scientist feels he has "proved" the "truth" of his hypothesis, which now assumes the status of a law - which is what the scientist originally set out to discover. The new law is made use of wherever the scientists feel it will be useful. The scientist is said to have added to man's knowledge of the universe: he has pushed back the frontiers of our ignorance.

The method of basing general statements on patterns which permeate accumulated data is called induction, and it is this that, traditional scientists claim, distinguishes science from other disciplines, which base their statements on such intangibles as prejudice, habit, speculation, tradition, authority, etc.

An alternative theoretical stance and methodological procedure arise from a negative answer to the following common sense question: Just because on one occasion event A occurs with event B, does it follow, logically and inexorably, that this will always be the case on all future occasions? Is there a mutual and incontrovertible dependence between A and B? It would not make a difference even if the researcher recorded a million occasions on which A and B were observed as concomitants. It is true that we would be psychologically disposed to expect A and B to co-occur. But to conclude from this that such a co-occurrence is an unalterable law of nature would be erroneous. Bryan Magee (1973, p.20),³ writing in the same vein, maintains that "... it cannot be established by logical argument that because all past futures have resembled past pasts.... all future futures will resemble future pasts."

The name Karl Raimund Popper (especially Popper (1959)⁴) has been closely associated with this alternative

3. Magee, B. (1973) Popper, Fontana

4. Popper, K.R. (1959) op.cit.

approach to the problem. He points out the diametric difference between the verification of a hypothesis on the one hand, and its falsification, on the other, and expresses his preference for the latter to the former.

"Now it is far from obvious, from a logical point of view, that we are justified in inferring universal statements from singular ones, no matter how numerous; for any conclusion drawn in this way may always turn out to be false: no matter how many instances of white swans we may have observed, this does not justify the conclusion that all swans are white." (1959,p.27).

Merely because we have recorded thousands of statements which report the sightings of white swans we can, inductively, proceed to claim that "All swans are white". On the other hand, the observation of only one black swan as opposed to a million white ones empowers us, deductively, to conclude that "Not all swans are white".

The inductive hypothesis after the observation of one black swan as opposed to a million white ones would be "Swans are either white or black". The essential difference between this hypothesis and the deductivist "Not all swans are white" is that the latter allows for the future appearance of a green, yellow or even a multi-coloured swan, whereas the former does not. This is a dogmatist versus pragmatist confrontation.

"All swans are white" is an empirical statement based on the observation of facts, an essentially limited (see Hempel below) number of facts. Although this claim cannot be proved, it is falsifiable, potentially: the scientist can go out into the field to look for a black

swan.⁵ He holds on to his hypothesis concerning the colour of swans only provisionally.

The narrow inductivist conception of scientific inquiry, whereby the scientist observes and records all the facts, and then analyses and classifies them, is altogether untenable. These first two steps, observing and recording all the facts, and then analysing and classifying them, are supposed to make no use of any guesses or hypothesis as regards the possible relationships among the facts. This seems to be a restriction imposed in order to rule out bias and the consequent jeopardisation of scientific objectivity. This attitude is not too far removed from that of the American structuralists in insisting on keeping the levels of linguistic analysis distinct. Hempel (1966, p.11)⁶ points out that the first step, namely the recording of all the facts, could never be carried out: a collection of all the facts would have to wait for the end of the world.

Popper also makes a clear distinction between what he calls the logic of the situation, and the implied methodology. On the logic-of-the-situation level, where we are concerned with the relationships between statements, a scientific hypothesis can be easily and conclusively falsified. The statement "All swans are white" is false if there is another statement "I admire black swans", and if the latter actually exist.

5. Later on in this thesis when counter-examples to a writer's conclusions are found, we shall call such examples "black swans".

6. Hempel, C. (1966) Philosophy of Natural Science. Englewood Cliffs, New Jersey, Prentice Hall.

The methodology level, on the other hand, presents a problem. It is possible to doubt a statement. There may be an error in the reported observation, we say. And, because the bird is black, we may therefore decide that it is not a swan. In other words, conclusive falsification is difficult to attain on the methodological level. So Popper suggests that the researcher keep these points in mind:

1. He must not evade the refutation of his hypothesis, whether by introducing ad hoc hypotheses or by ad hoc definitions, or by always refusing to accept the reliability of inconvenient experimental results;
2. On the contrary, he ought to formulate his theories clearly and unambiguously so that they can be exposed to refutation or modification in the light of new data.
3. Although the hunt for falsifying evidence is the Popperian rule of thumb, the scientist must not abandon his theories lightly. Doing so would betray an uncritical approach to tests. So a distinction has to be made between the readiness with which Popper accepts a refutation on the logic-of-the-situation level on the one hand, and the rigorous and highly critical attitude he adopts towards refutation on the methodological level.

The Popperian approach to research may be illustrated by considering the claim often made in some science textbooks, namely that "water boils at 100°C". It is useless to go about looking for confirming evidence, (which is what inductivists would do) because the truth of that statement cannot be proved. Instead, this

hypothesis can be tested by attempting to falsify it. And we soon find that water does not boil at 100°C in closed vessels.

In one stroke, we have destroyed the validity of the original hypothesis which our accumulated confirmations might have persuaded us to regard as a law of nature, a universal truth. (In reviewing a work of such monumental importance as A Comparative Study in Shona Phonetics, setting out to produce "black swans" may, at first sight, seem to be a particularly negative and foolhardy business. But since the review of the work on Shona phonetics is intended to place my own research into perspective and to point out those areas in which I believe I can make a contribution, this approach seems to be the most feasible and potentially productive. The merits of the work under review will not be glossed over.)

After falsifying the hypothesis in this way, Popper says, we should not reformulate it in order to accommodate the newly discovered discrepancy. Rather we should ask ourselves the all-important question "why does water not boil at 100°C in closed vessels?" In attempting to answer this question we are bound to come up with a more powerful hypothesis. The new hypothesis will explain both why water boils at 100°C in open vessels and why it does not in closed ones.

Our next step is to falsify the new hypothesis. In Popperian deductive science, this is the permanent next step. The traditional, inductive method of accumulating billions of confirming evidence poses no problems, and could increase the probability of the hypothesis being

accepted as the definitive truth. This would be misleading. The criterion of demarcation between science and non-science is, for Popper, falsifiability, not verifiability.

In linguistics the theory that seems to approximate to Popperian deductivism most closely is that associated with Noam Chomsky (1957, 1965, 1976)⁷ and with Chomsky and Halle (1968)⁸ and others of the same school. The transformational generative grammarians maintain that observation must be preceded by model-construction, i.e. hypothesization; for example the so-called L.A.D. (Language Acquisition Device). Since the number of possible sentences in any human language is infinite, theory must be based on a limited set of data from the language under investigation, but it (the theory) must both explain such data and be powerful enough to predict regularities which are likely to be found in hitherto unrecorded data from the same language.

"The critical problem for grammatical theory today is not a paucity of evidence but rather the inadequacy of present theories of language to account for masses of evidence that are hardly open to serious question ... the problem for one concerned with operational procedures is to develop tests that give the correct results and make relevant distinctions." (Chomsky, 1965, pp. 19-20).

-
7. Chomsky, N. (1957) Syntactic Structures, The Hague, Mouton
____ (1965) Aspects of the Theory of Syntax,
Cambridge, Mass: M.I.T. Press.
____ (1976) Reflections on Language, Temple Smith
8. Chomsky, N. and Halle M. (1968) The Sound Pattern of English, New York.

In one⁹ of the most recent extended expositions of Firthian linguistics it is suggested that the most viable theoretical and methodological stance would be a marriage between induction and deduction.

"Insistence on a basically inductive approach to analysis should not of course, be thought to preclude theoretical expectations and the need for predictive power in the classes and categories recognised in analysis. One can never be wholly inductive or exclusively deductive but it is possible at a given time to be more one than the other..." (p.7).

The cautiousness and pragmatism expressed by Mitchell here are shared by this author.

From Popper's point of view, that the history of science is as much one of revealed truths as of discarded hypotheses would seem to be a fair statement. Arthur Koestler expressed the same view in an address to the Conference of the PEN Worldwide Association of Writers. In his address¹⁰ The Vision that Links the Poet, the Painter and the Scientist, he said,

"... the same experimental data can in most cases be interpreted in more than one way - which is why the history of science echoes with as many venomous controversies as the history of literary criticism.... In fact the progress of science is strewn, like an ancient desert trail, with the bleached skeletons of discarded theories which once seemed to possess eternal life."

These upheavals have characterised linguistics no less than other fields of research.

-
9. Mitchell, T.F. (1976) Principles of Firthian Linguistics, Longmans.
 10. Koestler, A. (1976) The Vision that links the Poet, the Painter and the Scientist. In The Times newspaper, Wednesday, August 25th, 1976.

I am in agreement with what Mitchell (1976, p.7) says about the relationship between deduction and induction in the field of linguistics. The bone of contention seems to be whether, after formulating a hypothesis, further research should be aimed at proving or falsifying such a hypothesis. Either way, it seems that people from both schools are engaged in testing their hypothesis, the only difference being that those who set out to disprove their own hypotheses are more pragmatic in their approach.

Both the American Structuralists and the Transformationalists seem to be deductivists to the extent that they insist on the scientist formulating his theoretical expectations prior to investigating a language. Hockett's (1942)¹¹ article begins thus

"Linguistics is a classificatory science. The starting point of such a science is to define (1) the universe of discourse and (2) the criteria which are used in making the classifications. Selection and preliminary ordering of data determine the range of analysis; the choice of criteria fixes the level of analysis."

Two points emerge from this: that Hockett is one with the deductivists in (a) requiring that the field and scope of investigation be delimited; (b) in finding criteria for handling data. But he differs fundamentally from the transformational grammarians in restricting linguistic description to a mere classificatory science. Chomsky¹² demands that linguistic descriptions should have

11. Hockett, C.F. (1942) "A System of Descriptive Phonology" In Language, 18, pp.3-21. Also in Makkai (ed) Phonological Theory: Evolution and Current Practice, pp.99-112, 1972.

12. Chomsky, N. (1966) "The current scene in Linguistics: present directions." In Reibel and Schane (eds) Modern Studies in English.

explanatory power. He writes

"It is hardly open to question that the natural sciences are concerned precisely with the problem of explaining phenomena, and have little use for accurate description that is unrelated to problems of explanation."

This is his answer to those who criticised universal grammar on the grounds that it was misguided in seeking explanations, which the natural sciences allegedly did not do. Transformational-generative grammarians also require their hypotheses to have predictive power, that is, to account not only for the data under study, but also to explain that which may be collected later from the same language. This is not provided for in Hockett's remarks, nor, indeed, in those belonging to the same so-called taxonomist school.

Linguists who have applied the epithet "scientific" to their discipline have been at pains to emphasise the importance of the relationship between theory and data.

John Lyons (1962)¹³ maintains that

"The history of science is full of examples to support the opinion that the actual cannot be properly described, perhaps not even recognised, except in the framework of what has been previously envisaged as possible. At the same time, the sphere of what is thought as possible is being constantly revised under the impact of discoveries made in the description of actual languages."

The mutual dependence of theory and data is underlined.

Fry (1960)¹⁴ acknowledges the importance of this mutual

-
13. Lyons, J. (1962) "Phonemic and Non-phonemic Phonology: Some typological reflections." In Laver and Jones (eds) Phonetics in Linguistics: A book of Readings, Longmans, 1973.
 14. Fry, D.B. (1960) "Linguistic Theory and Experimental Research" In Jones and Laver, 1973 (eds.) op.cit.

dependence and cross-fertilization between theory and data, regretting, however, the bias towards theory at the expense of data in current linguistic research and writing.

"The general tendency has been to begin operations with a modicum of phonetic observation, construct a system of categories on the basis of such observations, and thereafter to devise theories, to modify them or discard them because of inter-relations within the theory rather than because of relations between theory and observations of behaviour."

I conclude this section by recalling a dichotomy which I referred to at the beginning: that the field of scientific enquiry falls into two sub-categories, the empirical and the non-empirical. Along with Lyons and Fry, I believe that linguistics will fall into the one category rather than the other depending on the extent of the cross-fertilization between theory and data. The farther away from data our theories of language become, the less empirical will our discipline be.

1.3 The American Structuralists and Phonemic Analysis

Contrary to what this uniform label might suggest, the linguists whose work is subsumed thereunder do not form a homogeneous group.

Bloch, Trager, Wells and Hockett typify Bloomfields' theories and practice in their work. From 1935 to 1965 Bernard Bloch was the editor of Language, the leading linguistics periodical in the United States. In collaboration

with Trager (1941)¹⁵ he published an analysis of the English phoneme system and the following year wrote their classic of structuralist linguistic analysis. They followed Bloomfield closely but gave more detailed directions concerning the establishment of phoneme inventories. And they also insisted on the allophones of the same phoneme meeting the requirement of phonetic similarity, something which other writers, e.g. Morris Swadesh (1934, 1935, 1936)¹⁶ had not found in Bloomfield. Bloch (1948)¹⁷ published an article dealing with the problem of segmentation, among other things.

The other school of thought is the one associated with K.L. Pike and his students. They are the least typical post-Bloomfieldians. The principal aim of Bloomfield and his successors was to arrive at an exact scientific description of linguistic structure using so-called discovery procedures. Pike is also interested in theory, but

15. Bloch and Trager, (1941) "The Syllabic Phonemes of English"

(1942) An Outline of Linguistic Analysis, Baltimore.

16. Swadesh, M. (1934) "The Phonemic Principle" Language; 10, pp.117-29. Also in Makkai (ed) op.cit., pp.32-39.

(1935) "Twaddell on Defining the Phoneme" In Makkai (ed.) op.cit., pp.41-44.

(1936) "The Phonemic Interpretation of Long Consonants" In Makkai (ed.) op.cit., pp.51-57.

17. Bloch, B. (1948) "A Set of Postulates for Phonemic Analysis" In Language, 24, pp.3-46.

his linguistic work has a direct practical purpose. Apart from being a Professor of Phonetics at Michigan, Pike is a missionary, leader and organiser of a number of courses in linguistics for missionaries at the Summer Institute of Linguistics. After their linguistic training these missionaries are sent to South and Central America, New Guinea and other places where they learn the local language, describe it scientifically, set up an alphabet, teach the natives to read and then translate the Bible into the language in question.

According to Pike, as also according to Daniel Jones, the phonological analysis of a language is meant to form the basis of an orthography which must be acceptable to the native speaker. In 1947¹⁸ Pike published Phonemics, significantly sub-titled a technique for reducing languages to writing. One of the main purposes for which research for A Comparative Study in Shona Phonetics was commissioned by the Government of Southern Rhodesia was to find a way of devising a uniform orthography for the dialects of Shona. This is why I think a comparison of the work of these American linguists with that of C.M. Doke on Shona could be informative.

This practical approach to language study probably explains why Pike was less dogmatic than the hard-core post-Bloomfieldians. He differs from them on a number of points. He is not afraid to include meaning in phonemic analysis; he does not insist on keeping grammar and phonology separate (on these two points, Doke and Fortune agree, in practice, with Pike) and wants not

18. Pike, K.L. (1947) Phonemics: a technique for reducing languages to writing. Ann Arbor. Univ. of Michigan Press

only to describe language, but also a number of other human activities according to the same methods.¹⁹ Pike was attacked by one of the hard-core post-Bloomfieldians for not keeping the levels separate.

In his review of Phonemics, Trager (1950)²⁰ observed

"If one has a good phonetic frame of observation, one gets phonemic solutions that hold together. But one cannot alter these solutions to fit one's preconceptions, one cannot mix up several levels of analysis and get anything but a mess."

In as far as Pike is concerned, language consists of three mutually dependent hierarchies: the phonological, the lexical and the grammatical, corresponding to three minimal units of analysis: the phoneme, the morpheme and the tagmeme. A derivation from the latter term, "Tagmemics", is used as a general designation of Pike's theory of language.

These two schools of post-Bloomfieldian American linguistic thought accepted certain common tenets regarding the structure of language. They both accepted that:

- (a) language has an underlying sound pattern beyond the overt stream of speech;
- (b) this system consists of a small number of segmental units, "phonemes";
- (c) this pattern could be discovered by applying to data a procedure consisting of the following criteria:
 - (i) free variation; (ii) complementary distribution;

19. Pike, K.L. (1967) Language in Relation to a Unified Theory of the Structure of Human Behaviour, Mouton & Co. The Hague.

20. Trager, G.L. (1950) Review of Pike's Phonemics in Language, 26, p.157.

34

(iii) contrast; (iv) phonetic similarity and (v) pattern symmetry/congruity.

The differences between these two schools arose from their attempts to define, and apply these criteria to data in order to discover the phoneme inventory of a language. Pike, as I have already said, used meaning in his definitions of these criteria, whereas the other linguists did not. These differences reflected their respective attitudes towards the current behaviourist philosophy of science. Pike made use of both the observable (phonetic phenomena) and the unobservable (meaning). As for Hockett (1942)²¹ and others,

"There must be no circularity; phonological analysis is assumed for grammatical analysis, and so must not assume any part of the latter. The line of demarcation between the two must be sharp."

Hockett's and his colleagues' suspicion of intangibles such as meaning, is responsible for the decline of interest in semantics during this period.

Pike defines the criteria (i) - (v) as follows: if two sounds occur in the same environment in a word without changing the meaning of the word, they are in free variation, and belong to the same phoneme. For example, in English the exploded [k^+] and the unexploded [k^-] in [$pæ k^+$] and [$pæ k^-$] "pack" are in free variation, since their occurrence does not entail a meaning difference between the two utterances. Complementary distribution or mutual exclusivity is defined as follows by Pike (1947, p.92)²²:

21. Hockett, C.F. (1942), op.cit.

22. Pike, K.L. (1947), op.cit.

35

"A set of three phonetically similar segments are mutually exclusive when the following can be said about them: 'The first segment occurs only in such and such positions, but the second and third segments never occur in these positions; the second segment occurs only in a different list of such and such positions, but the first and third segments never occur in this second list of positions; the third segment occurs only in positions not previously listed, whereas the first and second segments never occur in these positions.'"

Such a set of segments are allophones of the same phoneme.

Sounds which occur in the same environment in words which mean different things are said to be in contrastive distribution and belong to different phonemes.

Pattern symmetry is a criterion adduced when none of the criteria above can help the linguist decide whether or not some segments belong to the same phoneme. For example, if in the analysis of a language L we find that /p,b/ and /t,d/ are pairs of phonemes, but there are no minimal pairs or any other evidence to suggest that /k,g/ is another pair of phonemes (although they occur randomly in the words of the language) we can invoke this criterion and posit them as another pair of phonemes in the language. This conclusion would be based on the grounds that /k,g/ have the same phonetic relationship that holds between /t/ and /d/, and /p/ and /b/, namely the voiceless/voiced opposition as well as sharing the same place of articulation.

Phonetic similarity demands that segments which we classify in one phoneme must be related somehow in terms of their phonetic features, that is, articulatorily, auditorily and acoustically. For example, in English [ŋ] and [h] cannot be members of the same phoneme merely

because they are in complementary distribution:

#h -

- h#^{*}

- ŋ#

*#ŋ -

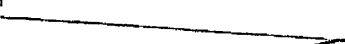

(where * = an unacceptable form, and # = word or syllable boundary).

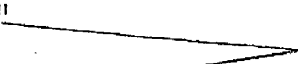
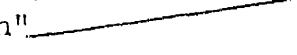
In terms of the deductive stance in scientific theory outlined above, every hypothesis about phenomena in the outside world makes specific claims and predictions about such phenomena. Pike's brand of structuralist phonemics is no exception. But it, like other theories, holds water to the extent that it gives an accurate account of the facts and no further. This theory has some fairly obvious merits. But there are some black swans.

Pike's definition of contrast above makes this claim: that phonemic differences correlate with meaning differences. This is not always the case as can be seen from the following examples from English:


"the apple", /ʔi - /

"the book", /ʔə - /, where the phonemic difference between /i/ and /ə/ does not entail a difference of meaning. Secondly, by claiming that phonemic differences correspond to meaning differences, Pike is also predicting that, conversely, meaning differences correspond to phonemic differences. But English has the following counter examples:

(a) (i) "way"  /wei/
(ii) "weigh" 

- (b) (i) "rain"  /reɪn/
 (ii) "reign" 

In other words, his theory does not cater for homophony. It cannot handle polysemy either. For example, Shona has the following example:

- ²³
 /godo/  "bone"
 "jealousy"

The major claim of this theory, as of that espoused by the other school of thought in American Structuralism, is that these criteria will give a phonemic analysis for any language. But these criteria do not solve the following problem from English:

- (i) /pɪt/ (ii) /bɪt/ (iii) /spɪt/

The examples (i) and (ii) give us the phonemes /p/ and /b/, by the criterion of contrastive distribution. However, the [p] in /spɪt/ resembles that in /pɪt/ in being voiceless; it resembles [b] in /bɪt/ in being unaspirated. Like both, it is a bilabial, pulmonic egressive plosive. To which phoneme, and using which criteria, are we to assign the [p] in /spɪt/? This is not a just a black swan for Pike's theory, but a failure which exposes the limitations of phonemic analysis of whatever orientation. Assigning this sound to the same phoneme as the [p] of /pɪn/, as is commonly done, is an act of faith.

The second theory, that championed by Bloch, Trager, Hockett and others, reflects more closely the behaviourist approach to scientific theory than the first. These linguists attempted to establish the phonemes of a language by what Bloch called "distributional analysis". All reference to meaning or other aspects of grammar was to be kept out of the whole business. Hockett^{23a} defines

^{23a} Hockett, C.F. (1942), op.cit. ²³ The two dots under a symbol indicate breathy-voice: I.P.A. Journal, June 1976

the criteria of contrast and complementary distribution as follows:

"If a and b occur in corresponding positions in utterances of different equivalence classes, they are members of different phonemes; if a and b... occur in mutually exclusive positions, they may be ... members of the same phoneme."

Thus, they operate with "utterance", rather than "word" or "phrase", in order to keep phonemics and grammar apart; with "equivalence classes" rather than "meaning", for the same reason. Items in mutually exclusive positions "may be" members of the same phoneme: the "may be" being a safety valve against the inclusion of segments such as English [ŋ] and [h] in one phoneme. Bloomfield (1933, p.139)²⁴ thought that the study of meaning should wait until the natural sciences had attained such sophistication as would enable them to attach a scientific label to every item in the lexicon, for example "salt" = NaCl.

The problem that immediately confronts us in applying these criteria to language data without reference to meaning or words and phrases is this: how do we tell whether we are dealing with contrast or free variation? And how do we define 'environment' (or 'positions') precisely without referring to units such as word, phrase or sentence? Because Bloch²⁵ defines free variation as phonetically different sounds with the same "environment". And contrast is when sounds have some, but not all, their

24. Bloomfield, L. (1933) Language, Henry Holt and Co., New York.

25. Bloch, B. (1948), op.cit.

"environments" in common. "Environment" remains essentially vague without reference to word, phrase or sentence.

Behaviourist empiricism demanded that only the observable be relied on in scientific research. Reference to meaning was to be excluded because that would be mentalistic. And the mind was the villain of the piece because of its "un-get-at-ability". The informant could be asked questions merely to elicit data, but his opinion about the data was unacceptable. The linguist's gospel comprises every word that proceeds from his informant's mouth - which cannot, by definition, be wrong; but as a matter of principle, whatever the informant volunteers about his language (as opposed to in it) must be assumed to be wrong.

The approach to language analysis advocated by these structuralists was strictly procedural. The linguist recorded and analysed data as follows:

5. Semantics
4. Syntax
3. Morphology
2. Phonemics
1. Phonetics



Each level presupposes a knowledge of the previous one and must not anticipate the one above. The linguist must start from the dark and work his way up, following the so-called "canons" of science as set out by Trager (1950)²⁶, among others : "... observe, formulate preliminary hypotheses, observe further, classify, test, formulate conclusions."

26. Trager, G.L. (1950), op.cit.

But when all is said and done, we have to ask why all these doubtlessly able, devoted and well-meaning linguists formulated their theories in this way and adopted their respective stances in relation to science. Their reasons immediately identify them with linguists who were operating in Africa during the same era. The Americans were, like, for example, C.M. Doke, trying to record and analyse some hitherto unwritten languages - the Americans the Ameri-Indian languages, and Doke, the Bantu languages of the African Sub-continent. They all espoused theories and methods which they thought best suited for their respective tasks.

The purpose of Doke's (1931, Preface)²⁷ work in Southern Rhodesia was

"the settlement of the language problems involving the unification of the dialects into literary form for official and educational purposes, and a standardization of a uniform orthography over the whole area."

The recommendations presented to the Legislative Assembly of Southern Rhodesia by Doke (1931)²⁸ do not concern us here.

A Comparative Study in Shona Phonetics is addressed to academics. In it, Doke reports his technical and scientific findings after two years' work in the field. No one who has not followed a course of training in Phonetics would need to be told that the work is technical: they would find it difficult to follow.

In conclusion, I think this is a good point at which to attempt a justification of the approach to the review adopted here: namely the prefacing of the review

27. Doke, C.M. (1931), op.cit.

28. Doke, C.M. (1931) Report on the Unification of the Shona Dialects, Stephen Austin and Sons Ltd.

with a brief exposition of the American phoneme theory and practice, assessed in terms of some tenets of scientific theory. It is not intended to suggest that Doke was, or should have been influenced by the Americans. The American phoneme theory is merely a vantage point from which to assess Doke's work since, as we have said already, he was engaged in similar research to that of the Americans. My standards in this appraisal are first, native speaker intuitions (i.e. my own), secondly, American phoneme theory and, lastly, the relevant tenets of scientific theory as outlined above.

Doke has been widely acclaimed, not undeservedly, by generations of Bantu scholars for devising a new theory of language suitable for Bantu. Thus Cole (1971)²⁹

"In The Phonetics of the Zulu Language, presented as a doctoral thesis in 1924, and published in 1926, C.M. Doke gave a preview of a new model for Bantu linguistic analysis and description... We have no information to suggest that Doke had any contact with or was influenced by the writings of Franz Boas, Edward Sapir, Leonard Bloomfield or other American linguists, who,... were and had been grappling with the problems presented in particular by Ameri-Indian languages."

The thrust of this quotation is that Doke's departure from common practice was in his grammatical description of the Bantu languages. Otherwise he was engaged in recording, analysing and classifying the sounds of strange languages, just as much as the Americans were. I see more scope for comparison on the level of phonetics and phonemics.

The strengths and shortcomings of the theories

29. Cole, D.T. (1971) "The History of African Linguistics to 1945" In T.A. Sebeok (ed.) Current Trends in Linguistics, Vol. 7, pp.1-29. Mouton.

within which African and American linguists were working in their respective areas can emerge in a comparative review such as this. Doke is thus being assessed by American standards as much as the Americans themselves are being assessed by Dokean standards: and both are being measured up against the relevant tenets of scientific theory, since they claim that their research practice was "scientific".

Fortune, especially Fortune (1968)³⁰, is a different matter altogether. He is torn between American Structuralist taxonomic linguistics and Dokean "Bantu grammar for Bantu languages." As H.A. Gleason, Jr. (1956)³¹ puts it, Fortune's phonetics "leans heavily on Doke's previous work" , while he (Fortune, 1968, p.5), quotes Bloch and Trager (1942)³², the hard-core post-Bloomfieldian behaviourists, with approval.

1.4 A Comparative Study in Shona Phonetics

From what we have said in (i) above about the nature of scientific theories and practice, we can draw the following conclusions about any study that claims to be scientific: first, that two things are necessary, namely suitable material for study, and the adoption of the methods of science. The latter can be summarised

30. Fortune, G. (1968) Shona Grammatical Constructions, unpublished student notes, Vols.I and II.

31. Gleason, H.A., Jr. (1956), Review of Fortune's An Analytical Grammar of Shona, Longmans, 1955, in Language, 32,3.

32. Bloch and Trager, (1942), op.cit.

as, observation (if we are inductivist), the devising of categories for the observed facts, the construction of hypotheses, the testing of these hypotheses by renewed observation, and, finally, the formulation of a theory accounting for the observed facts. The deductive school of thought would insist on a formulation of theoretical expectations prior to observation.

If this is taken as the definitive statement of the nature of scientific method, then I would submit that given the following circumstances,

- (a) that Doke was not a native speaker of Shona;
- (b) that he had one year in which to do field work;
- (c) that Shona has so many dialects and sub-dialects,
- (d) spread over a country three times the size of England;
- and,
- (e) that Shona, like most Bantu languages, has so much phonetic wealth,

very few people could have equalled, let alone surpassed, Doke's achievement. A Comparative Study in Shona Phonetics is the work of someone with special gifts. The achievement consists in Doke's accurate observation, recording, analysis and description of masses of data. There was no work on the Phonetics of "Shona" worth the name, that could have served as the starting point. In fact one of the aims of Doke's research was to unify the various dialects into one written language. The name of "Shona" is Doke's coinage.

A Comparative Study in Shona Phonetics is best discussed in two stages. There is the purely observational, analytical part of the work, that is, the phonetic description of the sounds of the dialects. Then there is the

phonological/morphological part, dealing with such processes as "nasalization", "velarization" and "vocalization".

In the first part of the study, Doke describes "the phonal elements of speech sounds", and, briefly, "tonetics", "dynamics" that is, "stress"/"prominence"; "duretics", that is, "the length or duration of phones or syllables", (1931, p.18). In this thesis, duration and length will not be used interchangeably as Doke seems to do. For me, duration is, like pitch, a phonetic feature, with a definite or absolute measurement on the appropriate scale. Length, on the other hand, is a more abstract, phonological feature, like the phoneme, made up of a potentially large number of allophones which are environmentally conditioned. I shall also use the term "speech sound" in a very specific way, such as that defined by O'Connor (1973, p.65)³³, "A speech sound is a perceptually distinct segment produced by a characteristic combination of organic actions." These clarifications are essential if ambiguity is to be avoided.

Doke divides the consonants of Central Shona into two groups: (a) Plain consonants, that is, those "composed of one phone element or a homorganic combination of elements", (a combination of homorganic elements perhaps?); (b) Velarised consonants, that is, "non-homorganic combinations which are due to the action of 'w', usually with bilabials."

(a) Plain Consonants. Doke points out the distinction made in Shona between implosives and explosives (1931, p.48)

33. O'Connor, J.D. (1973) Phonetics, Pelican original.

and describes the so-called whistling fricatives /ɕ/ and /ʑ/ (1931, pp.86, 87). Doke was qualified to make the remark (1931, pp.34-5) that Shona has more distinctive phone features than any other Bantu cluster, given his extensive and intensive research and publications on the languages of the Sub-continent,

"... the wealth of Bantu phonetics lies in the consonants, in their multiplicity of forms, their permutations and the limits and rules governing their combinations", (1931, p.36).

It is not necessary for me to reproduce here what Doke had to say about the articulatory and auditory features of individual sounds in Shona; except to mention that wherever possible and necessary, Doke's description of particular sounds was supplemented by

- (i) Kymograph records;
- (ii) Palatograms and diagrams showing places of articulation;
- (iii) Photographs of native speakers showing lip positions during the articulation of, for example, the lip rounded whistling fricatives singly and in combinations with other sounds.

I would, however, have preferred to use actual words as examples for palatography instead of the syllables used by Doke. One is more likely to get a more natural and faithful wipe using actual words, carefully selected so as to avoid articulatory overlapping in the palatal and alveolar regions of the mouth. Further, I would much rather have these palatograms and other instrumental evidence in the middle of the book instead of in the Appendices, where they are less accessible.

(b) Compound Consonants. After mentioning that the explosive consonants of Shona are produced at three points - the bilabial, alveolar/dental and velar, Doke (1931, p.37) says of 'p' that it is used in the "velarised compounds" "px", "pk" etc and in the affricate "pf". In my pronunciation of [-pxana], "crush", [p] and [x] are in sequence. To call [-px-] an example of velarization implies that the two consonants are produced simultaneously, which is not the case. In this thesis, secondary articulatory phenomena such as velarization and nasalization will be symbolised by diacritics, as is the usual practice in the literature. In my vocabulary I have the following items:

/mux^we/, "a crack", (n.3)

/mu^wɣe/, "on to" or "against a rock", (adv.)

In the pronunciation of the consonant 'cluster' at the beginning of the second syllable of each utterance, lip-rounding takes place at the same time as velar friction, accompanied by voicelessness in the first, and by breathy-voice in the second example. This gives me two sounds which are produced between the back of the tongue and the velum, with simultaneous lip-rounding, [x^w] and [ɣ^w]. This is an example of bilabialization of velar sounds - and not velarization. And since there are many other consonants which have this characteristic it seems reasonable to treat /x^w/ and /ɣ^w/ not as allophones of /x/ and /ɣ/, but as phonemes in their own right. They cannot be regarded as allophonic variants of /x, ɣ/ because they occur in identical environments, with their non-labialised counterparts: e.g. before /a/: /x^wānda/; "basket" vs. /xāɾānɪ/, "thread for sewing"

/x^wá x^wá/, "fat" (n.s) vs. /mùx^wáwù/, "greed".

before /o/: /ǰ^wòwà/ "mushroom" vs. /ǰ^wòní/ "deliberate malice"
(The matter is discussed at length in Chapter 5, pp 325-333).

In the language we also have the following minimal pair:

/-kárá/, "be greedy", "desire meat"

/-párá/, "scratch",

and therefore the phonemes /k/ and /p/. There is a word /-pkà^wà/, Zezuru for "crush". As we noted above, Doke calls "-pk-" an example of the velarization of /p/. For me, and for most people, "velarisation" is not synonymous with "combination with a velar". The latter is certainly the case in the complex -pk-. "Velarisation", "nasalization", "palatalization" etc. mean the superimposition of a secondary articulatory feature on to a primary. "Velarisation" in particular, means the raising of the back of the tongue towards the soft palate at the same time as articulatory activity essential for the production of an otherwise independent sound is taking place. And the latter sound occurs without such modification in the sound system of the language. If the raising of the tongue goes beyond a certain point we produce a fricative /x/ or /ǰ^w/; and if it is more radical still, we produce a plosive, /k/ or /g/. In such circumstances, we get clusters with the velar consonants as the second elements. In my dialect I can therefore find no phonetic plausibility for an analysis which treats [k] and [x], and [g] and [ǰ^w] as allophones of /w/, since I have the following minimal pairs:

/kù-wòr^wà/, "to decompose" - /w/

vs.

/kù-kòrà/, "to become fat" - /k/

vs.

/kù-xòrà/, "to earn, be paid" - /x/

/wàrà/, "temerity" - /w/

vs.

/-gàrà/, "sit" - /g/

vs.

/ràrà/, "dirt" - /r/

/-ròwà/, "bewitch" - /r/

vs.

/ǰ^wowà/, "mushroom" - /ǰ^w/

/ǰòní/, "deliberate malice" - /ǰ/

In my pronunciation [ǰ^wowà] and [ǰòwà] are free variants.

In his description of the voiced alveolar/dental plòsive, Doke (1931, p.42) says "The common 'd' of Shona is implosive, and the exploded 'd' occurs more commonly in compounds." This seems to suggest that Doke is positing one 'd' consonant with variants 'ǰ' occurring singly and more commonly, and 'd' occurring mainly in compounds. This is an unsatisfactory solution. Fortune (1968, pp.12-14) also makes the same mistake. Since Fortune extends the same interpretation to the bilabial plosives, I shall set out my arguments against this solution when I discuss Fortune (1968). Here I shall only mention the fact that the following data from Karanga (and the other dialects) does not support Doke's and Fortune's interpretation:

/-dàdà/, "be insolent" - /d/

vs.

/ḍàḍà/, "duck" (n.5) - /ḍ/

/ḍàkà/, irreconcilable hatred and desire for
revenge

vs.

/ḍàkà/, "mud"

/-ḍurà/, "confess"

vs.
/tʃaḍurà/, "it has been expensive"

Since /ḍ/ and /ḍ/ occur in the same phonetic environments, there is no sense in regarding one as an allophone of the other. Indeed, at a later stage (1931, p.51) Doke, in spite of what he says in the quotation above, admits that /ḍ/ and /ḍ/ are "semantically distinct"; in our terms, they are independent phonemes.

After describing the four nasal "consonants" in Shona, "m, n, ɲ, ŋ" (1931, pp.54-61), Doke adds,

"Apart from the four regular nasal consonants there is a fifth, the alveolar labialised 'nɸ', which, however, is not found apart from the voiced fricative 'z'."

Two points have to be made here. Firstly, all the nasals above are labialised when they are next to rounded vowels:

/munda/, a field

/-ɲura/, tear off meat from bone

/nom^ωé/, seven

/-ɲura/, drown

There is therefore no reason why 'nɸ' should be singled out and be called the "fifth nasal". Secondly, since /m, n, ɲ, ŋ/ are phonemes in the language, calling the 'nɸ' of "nɸ" the fifth nasal is likely to cause confusion by suggesting that it is also a phoneme. Doke does not seem to make any distinction between plain voice and breathy-voice: he

calls both /d/ and /d̥/, and /z/ voiced, although it is only /d̥/ which has plain - voice, while the others are breathy-voiced.

The widespread reference to "depressor" consonants in Shona linguistics literature *in recent years* (Fortune, 1968,³⁵ Fivaz, 1970,³⁶ Ladefoged, 1971³⁷), *springs from Doke's* realization that although in the orthography 'm' is combined with 'h' in such words as "mhanyá", "run", "mhuri", "family", in phonetic and phonological terms, this sequence of symbols represents one organic unit, the murmured or breathy-voiced [m̥]. To Doke's ears this sounded like a sequence of 'm' and 'h'. While detailing the distribution of the bilabial nasal /m/, Doke (1931, p.55) says it is found "before vowels, in mb, mv, succeeded by the voiced glottal fricative in mh, and in more or less velarised compounds varying from mŋ to mw."

Since /ŋ/ has been established as a phoneme (a "consonant" in Doke's terms), I cannot accept that /mŋ/ in some sub'dialects of Karanga /mŋàná/, "child", for example, is an instance of velarised /m/. It is a sequence of nasals, just as '-pk-' above is a sequence of orals, one bilabial, the other velar.

The five criticisms made of Doke's analysis and classification of the sounds of Shona so far:

(a) his treatment of /x/ and /ɣ/ as allophones or variants

35. Fortune, G. (1968), op.cit.

36. Fivaz, D. (1970) Shona Morphophonemics and Morphosyntax University of the Witwatersrand Press, Johannesburg.

37. Ladefoged, P. (1971) Preliminaries to Linguistic Phonetics, Chicago.

- of /w/;
- (b) his classification of "-pk-" as velarised "p";
 - (c) his description of /d̥/ as an allophone of /ḍ/, while contradictorily admitting later that the two are semantically distinct;
 - (d) his positing of the labialised "n" in "n₂" as the fifth nasal, thereby, indirectly giving it phonemic status, and finally,
 - (e) his failure to realise that /m/ is one organic unit rather than a sequence of 'm' and 'h',

spring from one fundamental weakness of Doke's phonetics as presented in A Comparative Study in Shona Phonetics: his failure to back up largely accurate observation with an explicitly articulated phoneme theory such as that of his American and British contemporaries.

He merely divides the sounds of Shona into the two broad traditional categories - "consonants" and "vowels". Having recognised "the multiplicity of forms" of the Shona sound system from the beginning (Doke, 1931, p.36), Doke should have found or devised some categories capable of bringing some order to this otherwise tantalising multiplicity. And it is my firm belief that the phoneme theory, modified to be accommodated within the Dokean Bantu grammar for Bantu languages if need be, would have done just that. This shortcoming on Doke's part was pointed out by H.A. Gleason, Jr. (1956)³⁸ in his review of Doke's Southern Bantu Languages:

"... the Doke pattern has serious limitations... the first is its weak development of phonemics. The phoneme concept is not explicit anywhere in the book..."

38. Gleason, H.A. Jr. (1956) Review of Doke's Southern Bantu Languages, in Language, 32,3.

As a matter of fact, E. Fischer-Jørgensen (1975, p.71)³⁹ counts H.A. Gleason, Jr. among the more dogmatic structuralists. The point I wish to establish, however, is that Doke's exclusion of the phoneme theory from his work on Shona seems to have been deliberate. If he did not know about the work of Bloomfield and his successors when he wrote A Comparative Study in Shona Phonetics in 1931, he most surely did when he wrote Southern Bantu Languages in the 1950's in which, according to Gleason, the phoneme concept exerted no influence. The question is, why? My criticisms above show that Doke paid a big price by not benefitting from his intellectual milieu.

In the same review, Gleason gives Doke credit for breaking away from the European tradition and for devising categories more suitable for Bantu languages. This does not seem to be the case in Doke's phonetic work. He seems, in that field, to have starved himself of the best achievements of the American and European traditions. Avoiding all reference to the phoneme concept, he deals instead with "consonants" and "vowels" in both the phonetic and the phonological parts of the book without any indication as to whether these labels are articulatory or phonological. Other linguists, for example Pike (1944)⁴⁰ would make the following crucial distinctions: vocoids and contoids are phonetic labels, referring exclusively to the articulatory, auditory and acoustic features of segments; while vowel phoneme and consonant phoneme are phonological labels,

39. Fischer-Jørgensen, E. (1975) Trends in Phonological Theory: a historical Introduction, Copenhagen, Akademik.

40. Pike, K.L. (1944) Phonetics, Ann Arbor, University of Michigan Press.

referring to the role played by the respective segment types in the structure of the larger units of linguistic analysis, e.g. the syllable or morpheme. Since Shona (all the dialects) has a CV syllable structure, the most primitive and simplest, these categories would come in handy in a phonemic analysis.

New categories and terminology devised to handle phenomena peculiar to African languages are, of course, welcome. But using terms such as "velarisation", already well established in European and American linguistic theory, giving them local specialised definitions (or no definitions at all), can only result in confusion and unnecessary complication. Since in recent years the sharing of ideas between Euro-American and African linguists has reached unprecedented proportions - many volumes of the series Current Trends in Linguistics⁴¹ are devoted to language groups in Africa - it would be useful if researchers did not give local overtones to terminology which is already widely used in general linguistic theory.

The most important suprasegmental feature in all the dialects of Shona is tone. As we have seen, Doke's work on the segmentals was supplemented by field instruments. But he does not seem to have had any instruments or technique for handling tonal and pitch phenomena. This explains the rather unsatisfactory tone marking scattered in the book. And I must agree with Gleason (1956)⁴² who says that

"... Doke's tone markings are certainly the least satisfactory part of his phonetics.

41. Sebeok, T.A. (ed.) (1971) Current Trends in Linguistics, Mouton.

42. Gleason, H.A. (1956), *op.cit.*

24

This is true not only for the linguist but also for the native speaker... Lacking a technique for handling pitch phenomena, Doke must omit them or pass over them incidentally and impressionistically, to the detriment of his morphology... there are many places where an overtly phonemic analysis would make for greater clarity, greater completeness, or both."

Although Gleason was writing about Doke's work on Bantu languages generally, these remarks are particularly apposite to Doke's performance in A Comparative Study in Shona Phonetics.

Doke is correct to say that stress is *not used for* emphasis in Shona. But I must disagree with everything else he has to say on the matter (Doke, 1931,p.205):

"Stress is the word builder. The stressed syllable gathers around itself the unstressed syllables, and unites them into one word."

He writes this although in the immediately preceding paragraph he has given examples of four Shona sentences, showing how this language has to use different word order and to introduce fresh lexical items to translate four English sentences which merely differ in the placement of emphatic stress. Whatever else stress does in Shona, it most certainly does not build up words in the manner suggested by Doke.

Incidentally, the rendition of the English sentences (1931,pp.204-5) given by Doke(1931,p.205) will

amaze not only every native speaker of Shona, but all those who otherwise correctly acclaim Doke for pioneering Bantu grammar for Bantu languages:

English

Shona

- | | |
|-----------------------------|--|
| (i) 'We want you to speak | (i) isú tinodá kuti utáure |
| (ii) We 'want you to speak | (ii) (Kudá tinodá kuti utáure.
(tinodá zokwadi kuti utáure. |
| (iii) We want 'you to speak | (iii) tinodá kuti iwé utáure |
| (iv) We want you to 'speak | (iv) tinodá kuti utáure nokutáura |

Shona people do not speak like this. These sentences are not Shona, let alone Karanga.

The processes which Doke calls "nasalization" (1931, pp.62-71) and "vocalization" (1931, pp.125-135), that is, the phonological part of the book, will not be discussed here. Fortune (1955, 1968)⁴³ discusses similar phenomena but does not depart too radically from Doke. These processes are discussed at great length in chapter 6 (pp.335-384) where I am proposing an alternative analysis to the taxonomic approach adopted by both Doke and Fortune.

Here, I think I ought to defend Doke against H.A. Gleason's other criticisms set out in his review of Doke's Southern Bantu Languages. Gleason notes that no distinction is made between phonology and morphology by Doke.

"Doke's model also fails to make a distinction between phonologic and morphologic levels of structure."

But, remembering what he has already said about the lack of any phonemic awareness on Doke's part, Gleason qualifies this remark:

"It is open to question whether in the absence of rigorous phonemics it is worthwhile to insist on thorough separation of levels."

43. Fortune, G. (1955) An Analytical Grammar of Shona, Longmans.
____ (1968), op.cit.

We are not told in what way, if Doke had utilised the phonemic principle, his description of Bantu would have benefitted by thorough "separation of levels."

The American Structuralists' insistence on this kind of separation of levels is based on two views of language structure ; one is that language is actually organised in levels ; the other is that these levels are levels of analysis, that is, they are creations of the linguist designed to facilitate the reporting and description of language data. Otherwise, according to this latter view, the object of study, language, is unitary. If Gleason is criticising Doke for not holding the first of these two views, then Doke must be congratulated for not falling into that post-Bloomfieldian trap. As Trager (1950) has said in his review of Pike's Phonemics, the post-Bloomfieldian phonemicists believed that "... one cannot mix-up several levels of analysis and get anything but a mess."

If Doke's A Comparative Study in Shona Phonetics, and the section on phonology-cum-morphology in particular, is a mess (and this would be an exaggeration), it is so for reasons other than that he mixed-up levels of analysis. I can only regret that Doke seems to have been too theory-shy to declare his position on these matters.

Conclusion

What has been said about A Comparative Study in Shona Phonetics suggests that we include Doke among linguists who, when doing phonetic analysis, ask of every

feature, 'Can it be related to the grammar?' rather than among those who ask, 'How can we keep the grammar out?'

The following are some of the points that can be made about C.M. Doke's work on Shona in relation to what we have said about

(a) scientific theories:

- (i) he made accurate observations and records of data in the field;
- (ii) the majority of his conclusions are based on these data;
- (iii) there are instances where he fails to interpret or classify facts because he has no well-defined theoretical framework;
- (iv) the juxtaposition of his claim about the function of stress in Shona with data that contradicts this claim is the most disturbing "black swan" in this part of the book.

(b) American phoneme theory:

- (i) he used meaning in the phonetic analysis and classification;
- (ii) he made no distinction between levels of analysis;
- (iii) this did not lead to a mess, contrary to Trager(1950
- (iv) the unsatisfactoriness of his description of certain processes hinges on his fundamental failure to declare his theoretical orientation in A Comparative Study and to stick to it.

Doke lies in the twilight zone between the dogmatic structuralists, and the pragmatic ones led by Pike.

Consequently he does not seem to be exclusively inductive

or deductive.

1.4.1. An Analytical Grammar of Shona.⁴⁴

Although, as Gleason⁴⁵ says in his review of Fortune (1955)⁴⁴ Fortune's phonetics and phonology chapters "lean heavily on Doke's previous work", there are a number of points made by Fortune which are worth commenting on. The two linguists' approach is, however, generally the same in that Fortune, leaning heavily on Doke, does not work within a well-developed phoneme theory. He uses the terms 'consonants' and 'vowels' in much the same way that Doke does. 'Velarization' is also used to mean 'combination with a velar'.

It is important to note that Fortune's grammar is based mainly on the Zezuru dialect, although he frequently refers to "Central Shona", a term used by C.M. Doke (1931)⁴⁶ to refer to the Korekore, Manyika, Karanga and Zezuru dialects together.

On his "chart of plain consonants" Fortune (1955, p.6) has a number of symbols in brackets, for example, "^t~~t~~, d, ~~x~~, ~~y~~" with a footnote reading "consonants in brackets are found only in combination with other sounds." The terms 'consonant' and 'sound' are being used synonymously. This, as I have already said, is rather confusing and will be avoided in this thesis. The semi-

44. Fortune, G. (1955), op.cit.

45. Gleason, H.A. Jr., (1956), op.cit.

46. Doke, C.M. (1931), op.cit.

independent status given to these bracketed 'sounds' recalls Doke's treatment of the so-called labialised fifth nasal 'ɲ' in /ɲz/. The same objections may be raised here. For example, why does Fortune not put [s] in brackets since it has lip-rounding in /-súmúkǎ/, "stand up"? The phonetic relationship between [t] in /tʃ/, for example, and [t] in /-támǎ/, "play" is obscured if we express it in Fortune's terms.

Working within a theory such as the phoneme theory, we can establish a phoneme /t/ and say that the two types of [t] above are allophones of this phoneme. Once this is established there will be no need to include the diacriticised [t] on the chart - even in brackets - which, I presume, is a list of the distinctive segments of the language and not of the numerous conditional variants.

Finally, it does not make sense to say that the "consonants in brackets are found only in combination with other sounds" since all the segments of a language must combine "with other sounds" to form words and phrases: a segment has no worth except in its relation to other segments in the same system.

Fortune also divides the consonants of Central Shona into two sets - the plain and the compound. The plain consonants are "those of single articulation", and the compound ones "those with double articulation" or whose production takes place "at more than one point" (Fortune 1955, pp.6, 10). For example /p,b/ are plain, and /ɕ,d/ are compound consonants. The former are plain because they have bilabial closure only, whereas /ɕ/, for

example, has friction in the alveolar area as well as lip-rounding. By the same token, one would expect /w/ to be called a compound consonant or semi-vowel. This, however, is not the case. In fact /w/ is not included on the two charts of plain or compound consonants. This is my first criticism of the plain vs. compound consonant dichotomy: that it is not consistently made.

The second criticism is more serious. This dichotomy has no phonetic plausibility. For example, it is simply not true to say that the 'consonant' /b/ has "single articulation". If, in addition to the bilabial closure, there were no velic closure to cut off nasal resonance and no vocal fold vibration and no egressive air stream mechanism, there would be no /b/. Fortune should give us some reasons for his apparent belief that lip-rounding for /ɣ/ is more important than the velic closure for /p/ or /b/, for example. All consonants have activity at more than one point, in the speech tract.

The sound [ɣ], in /-ɣa/, "become", is not a fricative for me and other Karanga speakers, contrary to Fortune (1955, p.8), but a frictionless continuant.

In his recent writings, for example Shona Grammatical Constructions (1968, p.16)⁴⁷, Fortune no longer regards the syllable onset in "mhanya", /-ma^hɲa/, "run", as a sequence of m and h phonetically, which he did in this grammar (1955, p.12), no doubt following Doke. He treats it as one breathy-voiced phoneme.

Doke's rather clumsy handling of tonal phenomena in Shona was, retrospectively, a gallant attempt to explain

47. Fortune, G. (1968), op.cit.

what must have been a puzzling aspect of the language structure. Fortune, on the other hand, not only omits all mention of tone, but he does not mark it on any one of the thousands of examples he gives. And yet H.A. Gleason, Jr. (1956)⁴⁸ was convinced in his review of this book that

"Fortune wrote on Shona not as a linguist who has done intensive short-term field work with a grammar as his immediate and perhaps final goal, but as one who speaks the language and has used it extensively... Fortune knows Shona."

This is an extravagant claim to make for Fortune. It is inconceivable that anyone can write even a tolerable grammar - analytical or not - of a tone language without mentioning, let alone positing some rules accounting for, the tonal processes that are an integral part of such a language.

Unlike Doke, Fortune (1955) divides the PHONETICS section of his book into Phonetics and Phonology. My comments above were directed at what Fortune says in the Phonetics section. In the Phonology section Fortune deals with processes involving the commutation of phonetic features leading to the formation of new words belonging to various classes. For example

verbs -----> nouns

nouns (class x) -----> nouns (class y)

adjectives ----> nouns

Of the processes which Fortune discusses under the heading of phonology, I wish to draw attention to the following:

48. Gleason, H.A. Jr. (1956), op.cit.

Voicing (1955, p.41) - defined as "the voicing of a non-voiced sound". We are told that "p, t, k, t̥, ts and, rarely pf become voiced when they occur initially in the stems of class 5 nouns or adjectives."

e.g. bángá	'knife	the plurals	màpángá
démó	'axe'		màtemó
dzindà	'chief's son'		màtfindà
dzát̥d̥zát̥sát̥sá	'spider'		màtsát̥sát̥sá
byènbvèni	'baboon'		màpfèni

I should point out that in the case of the first two examples, it is not enough to say that [p] and [t] have become voiced: they have also become imploded, in the derivation of their respective class 5 nouns. Fortune does not give the stems from which these class 5 nouns are derived. He, however, draws attention to the plurals and says the voiceless consonants become voiced. Doke (1931, pp.126-130) gives the stems. For example, for

bángá - pángá
dzindà - tfindà, and so on.

The '-' before the voiceless consonant-commencing stems indicates that a prefix has been omitted, in these cases /mà-/. This is the class 6 noun prefix, i.e. the numerical plural of class 5 nouns.

When this process ("voicing" for Fortune and "vocalization" for Doke) is described as one whereby class 5 nouns or adjectives are derived, the implication is, therefore, that this is a process whereby singulars (class 5) are derived from plurals (class 6). This raises a theoretical as well as a practical problem.

The theoretical problem is whether we should set up the plural stems of words as the basic form from which

other items in the formal scatter are derived. Fortune and Doke appear to have decided that the plural stems should be basic. This is both counter-intuitive for me and contrary to common practice in linguistics. The usual thing is to have a rule whereby plurals are formed from singulars, and not vice versa. If we reverse the processes described by Doke and Fortune, that is, instead of "vocalizing" or "voicing" to form singulars from plural stems, we devocalize to derive plurals from singular basic forms, and I am suggesting that we do:

e.g. instead of

-páŋgá (pl.) ----> báŋgá
 -tʃíndà(pl.) ----> dʒíndà

we have

báŋgá (sing.) ----> -páŋgá
 dʒíndà(sing.) ----> -tʃíndà

we are not going to have a simpler, or even a more complicated rule, since both formulations involve the commutation of the same number of phonetic features. But what we shall gain is the satisfaction of native speaker intuitions.

Dr. N.C. Dembetembe⁴⁹ has drawn my attention to the formal scatters:

-páŋgá:/màpáŋgá/'knives'
 /rùpáŋgá/thin knife
 /kàpáŋgá/short knife
 -tʃíndà:/ùtʃíndà/being a chief's representative
 /mùtʃíndà/a chief's representative
 /màtʃíndà/chief's representative

49. Dr. Dembetembe, N.C. of Department of African Languages, Un. of Rhodesia - Personal communication.

and suggested that the grammar would seem to be made simpler by positing the forms which commence with voiceless segments as basic forms. This, however, does not seem to be the case, although the relatedness of the forms in each formal scatter is brought out more clearly thereby.

The second criticism relates to the practical problem posed by Fortune's and Doke's solution. The strongest argument against setting up these voiceless forms as basic is that being plural stems, they are not the ones entered into the lexicon. I propose that we have the singular forms in the lexicon of Shona, leaving out the plurals, which can be derived, from the singulars after lexical insertion, by rules of the phonological component of transformational generative grammar.

"Assimilation" (1955, p.26) is defined by Fortune as a process whereby "a sound is made similar to a neighbouring sound." He discusses a particular type of assimilation called "bilabialization", i.e. when "the nasal consonant of the prefix of classes 9 and 10 is bilabialised when joined to stems commencing in a bilabial or labio-dental consonant" (Fortune, 1955, p.36):

- A. N+~~w~~ona, "see" ---> mboni, "pupil of eye";
N+puka ---> mhuka, "animal";
N+-berek'a "beget" ---> mberek'i "parent";
N+-vuta "blow" ---> mvuto "bellows".

Another, apparently different, type of assimilation discussed is called "supradentalization". "The nasal consonant of the prefixes of classes 9 and 10 is supradentalised when joined to stems commencing in supradentals or alveolars" (Fortune, 1955, p.36):

- B. N+-tema "cut" ----> nhemo "chisel"
 N+-diki "small" --> ndiki "small"
 N+-refu "long" ----> ndefu "long"

Yet another process is discussed, called "frication" - "the turning of a non-fricative consonant into a fricative" (Fortune, 1955, p.42). This is, apparently, not another type of assimilation. "The consonants p, t and k, when occurring as initial phones of stems of classes 9 and 10, are changed into the voiced glottal fricative h.":

- C. ru-puka "thin animal" ----> mhuka "animal"
 -tem "cut" ----> nhemo "chisel"
 hu-kama "relationship" ----> hama "relative"

The underlined example in A above is essentially the same as the one in C, also underlined. This is confirmed by the fact that the same word "mhuka", "animal", is derived from the examples. But the process involved in this derivation is called "bilabialization" in A, a type of assimilation, while it is called "frication" in C. Similarly, in B above, (a type of assimilation called "supradentalization") we are given the example

N+-tema ----> nhemo

as an example of assimilation, while the same example is given in C as an example of frication. Also in A, we are given, as an example of assimilative bilabialization,

N+-bereka ----> mbereki

while in B,

N+-diki ----> ndiki

is given as an example of supradentalization. These two

sets of examples demonstrate, for me, the same process, since in each case an implosive + homorganic nasal becomes the explosive preceded by a homorganic nasal.

The criticism I am making (and will elaborate in chapter 6 below, (pp.335-84)) is a fairly serious one: Fortune misses many opportunities for making useful generalizations which would simplify his grammar considerably. As Matthews (1974, p.119)⁵⁰ says,

"The aim of morphological analysis, as of all linguistic analysis in general, is to bring out the relationships which are systematic in the given language."

In his review of An Analytical Grammar of Shona, Gleason (1956)⁵¹ expressed the view that the definitions which preface almost every section in the book were a "ritual".

"The reader should regard his definitions as a ritual and disregard them. Nothing is changed because they are wholly external to the fabric of his description."

This is rather difficult to do since the processes so defined lead to what seem to be false dichotomies among what are essentially the same phenomena.

"In some cases", continues Gleason, "it is clear that a more scientific method would have avoided serious defects". I could not agree more.

"But it might also have destroyed some of the real values of the book if it had been used with the same artless rigidity that some scientific linguists have cultivated."

There is no doubt that the ritualistic definitions

50. Matthews, P.H. (1974) Morphology: An introduction to the theory of word-structure, London, Cambridge Univ. Press.

51. Gleason, H.A. Jr. (1956), op.cit.

in Fortune's grammar contribute something to the artistry which Gleason thus acclaims. If being artistic can be reconciled with being scientific, there would be no objection from me. But Fortune gives three such definitions - of bilabialization, supradentalization and frication - for what is essentially the same process without even indicating that he is aware of the underlying common system.

Finally, yet another type of assimilation, "supradentalization and alveolarization" (1955, p.37) is said to take place "in the formation of short causatives". There is no indication as to how short such causatives are but all the examples given by Fortune are disyllabic:

-muka "rise" +yâ ----> -mutsa "rouse"
 -rira "cry" +yâ ----> -ridza "sound"
 -ipa "be bad" +yâ ----> -iṣa "make bad"
 -reba "be long" +yâ ----> -redza "lengthen"

The formation of causatives, whether or not we can call it "supradentalization and alveolarization", occurs with examples of three syllables:

/-rúmúra/"wean" -----> /rúmúdzà/"cause to wean"
 /-bèrèkà/"carry on back" ----> /bèrètsà/"make to carry on back"

and with examples of four syllables:

/-nánájírà/"move slowly" ----> /nánájídžà/"cause to walk slowly"
 /-ṣúngúrúkà/"be frustrated" --> /-ṣúngúrúdzà/"cause frustration" or "disappointment"
 /-pxítítíkà/ rise (of smoke or dust) ----> /-pxítítídžà/"cause smoke to rise".

1.4.2 Shona Grammatical Constructions.⁵²

Shona Grammatical Constructions was first circulated by the Department of African Languages, University College of Rhodesia, in 1968-9. The Phonetics and Phonology parts of this work represent a fundamental departure from An Analytical Grammar of Shona in one important respect: there is a deliberate attempt to develop a phoneme theory and to work within such a theory. Although Doke's shadow is still all too obvious (Doke (1931) is quoted with approval four times within the first thirty pages), the section devoted to the segmental sound units of Central Shona is phonemically orientated.

This is to be commended, not simply because it is a phonemic analysis, but because those who read it will find it easier to agree or disagree with the author within this particular theory of language structure. This is conducive to the development of a brand of phonemics suitable to the language in question and to the language group as a whole.

There are two sections on the first page, headed Morphemes (1.1) and Phonemes (1.2) respectively. The main thrust of his definitions of these units is that whereas morphemes are meaningful, phonemes and syllables, of which morphemes are made, have no meaning. I have nothing to add to Fortune's characterization of the morpheme. But while it is the case, for Shona at least, that a

52. Fortune, G. (1968) Shona Grammatical Constructions. Unpublished student notes. Dept. of African Languages Univ. of Rhodesia.

phoneme per se has no meaning, I think we ought to mention the important fact that it has the capability of bringing about a change of meaning. For example, if the phoneme /m/ in

/mũ̀ũ̀/, "person"

is replaced by the phoneme /t/, we get an entirely new word

/tũ̀ũ̀/, "things".

Section 1.5.1 is devoted to Egressive and ingressive air streams (1968, p.5). Fortune writes,

"In Shona most consonants are egressive but the voiced stops are ingressive or implosive, i.e. released inwardly on an ingressive air-stream."

Since at this stage he has not mentioned depressor, or, (more appropriately at this phonetic level), murmured/breathy-voiced stops, the reader is likely to be confused by that statement. In Shona we make a distinction between plain-voiced and breathy-voiced sounds. It is to the former that Fortune is referring, and to the two implosives (he thinks there are three) in particular, /ɓ,ɗ/.

Section 1.5.2.1 (1968, pp.5-6) is on Voice, voicelessness and murmur.

"Voice is always produced at a certain pitch depending, mainly, on the tension at which the vocal cords are held while vibrating."

Fortune does not say what other factors there are which help determine voice pitch. But in addition to their tension, their length, thickness and the amount of lung air passing between them, are all contributory factors and have a bearing on the fundamental frequency of the sound wave produced by the vocal folds. The situation may be represented schematically as follows:

thick vocal folds	}	the same length	}	lower pitch
		the same tension		
thin vocal folds		the same lung air pressure		higher pitch

Ladefoged (1971,p.7)⁵³ agrees with me in thinking that more than one factor plays an essential role in determining pitch:

"The rate of vibration during a voiced sound depends on two factors: the tension of the vocal cords, and the pressure drop across them."

In the same section (1968, p.6) Fortune writes,

"In this study, sounds of which murmur or 'breathy voice' is a component, are called voiced depressor."

There are two points to be made here. The first is terminological: Shona does not have voiceless depressor consonants. So Fortune's epithet 'voiced' before 'depressor' is redundant. The second point is related to levels of analysis and their appropriate terminology.

These consonants are called "depressor" because they have the general tendency to lower syllable pitch. It follows that since the syllable is a phonological unit, the term depressor cannot be used to describe individual segments in isolation. Murmur or breathy-voice, on the other hand, describes the sound segment and is therefore the appropriate phonetic label. This is one instance in which we can insist on a separation of levels, and terminology and achieve something positive: clarity of description.

53. Ladefoged, P. (1971, op.cit.

(1975, p.224). "Altering the tension of the vocal cords is the normal way of producing most pitch variations that occur in speech."

Fortune groups the nasal phonemes and the plosives in one category, the stops.⁵⁴ He, however, characterises stops in section 1.5.2.3 (1968, p.7) as follows,

"The articulators may come into complete contact and thus stop the air which is then released,"

something which does not happen when nasals are pronounced. He then makes a distinction between "radical" or "aspirated" stops on the one hand, and "unaspirated" stops on the other. In section 1.6.4 (1968, p.14), The nasal stops, Fortune does not say anything about the manner of release of these so-called stops, nor does he say anything about aspiration.

My point is, the grouping of plosives and nasals in one segment category stops has no obvious phonetic motivation. In section 1.6.10 where Fortune discusses The Consonant Clusters (1968, p.19), affricates are defined as "consisting of stop and spirant, both voiceless and voiced depressor, viz. /pf, ts, tʃ, bʏ, dʒ, dʒ and dʒ /". Fortune's "nasal stops" are conspicuous by their absence from this list of Shona affricates, in spite of the fact that some of them combine with identical "spirants": /mv ~ bv, nz ~ dz, mʒ ~ dʒ /. If both nasals and plosives are stops, why does Fortune not list all their combinations

54. I am aware of the fact that these two classes of segments are grouped together by Generative Phonologists, and in particular by Chomsky and Halle in The Sound Pattern of English, where both are said to be [-continuant]. But since at this stage we are concerned to identify the phonetic characteristics of the respective sound classes, I think my objections are in order.

with spirants as affricates? He instead places the nasal + spirant clusters in a different category called "nasal spirant clusters, voiced depressor" (1968, p.20). There is no feed-back between Fortune's phonetics and phonology.

In spite of the articulatory and phonotactic evidence against the classification of nasals and plosives together as stops, Fortune (1968, p.24) characterises the nasals as stops on his consonant charts. And the fricatives which combine with the nasals - see his Clustered Consonantal Chart - are called "nasal spirants". The most these fricatives could be is nasalised. And I have x-ray records in the section on naso-oral consonants in chapter 5 (pp.198-333) where these fricatives share none of the nasality of the adjacent nasals.

Fortune defines the fricatives by saying that they "differ from stops in that the flow of air, though impeded, is continuous (for nasals the air is channelled via the nose and flows continuously!)". The sentence in brackets comes rather discordantly after the classification of nasals as stops. But he continues,

"Another type of continuant, to which Ladefoged has given the useful term approximant, involves an articulation in which the articulators approach each other yet without producing a turbulent air stream."

The examples given are the semi-vowels [w,j] and the bilabial frictionless continuant [β].⁵⁵ In the production

55. In this thesis I shall use the symbol [ν] for the labio-dental frictionless continuant, which I consistently use and hear from my Karanga interlocutors, instead of the bilabial frictionless continuant [β].

of vowel sounds the articulators also approach each other in the manner described by Fortune. There are some phoneticians who use the term "approximant" to refer to all frictionless continuants, which includes vowels.

But, as I shall elaborate in the next chapter, I propose to use the term as follows: both frictionless and fricative continuants are approximants, in the sense that the articulators approach each other without forming a closure. The difference between them is this: for fricative continuants, the articulators attain a narrow degree of approximation, hence the friction; whereas for frictionless continuants, the articulators, though approaching each other, retain a wide degree of approximation, causing no friction. In both cases the stricture is maintained for an appreciable stretch of time.

This way of viewing the sounds seems to be backed up by acoustic evidence since on the spectrograms there is continuous energy on the spectrum for both sound types, the difference being that whereas the frictionless continuants have their energy arranged in formants, the fricatives have a random spectrum, especially in the higher frequencies.

Semi-vowels are very transitory and should not, strictly speaking, be included among continuants. Thus Gimson (1962) "A semi-vowel is a rapid vocalic glide on to a syllabic sound of greater steady duration." And, "Compared with /r, l/, the steady state of the semi-vowels is very short, e.g. of the order of 30 msec, "(1962, p.213).⁵⁶

56. Gimson, A.C. (1962) An Introduction to the pronunciation of English, Edward Arnold, 2nd Edn. 1970.

Fortune's failure to distinguish between semi-vowels and continuants on the grounds of duration leads to his inclusion of flaps among the continuants, "A final group of continuants includes the lateral, flapped and rolled consonants," (1968, p.19).

Section 1.5.4 is on Double Articulation. This is defined by Fortune as follows, "It is possible for two articulations to take place at different places at the same time." In this exercise, for an articulation to be characterised as "double", it must meet one pre-requisite: there must be two strictures of equal importance.⁵⁷ For example, in the pronunciation of a click, there are two closures at different points in the oral cavity. The examples given by Fortune do not meet this requirement. He mentions the whistling fricatives of Shona, /ɬ, ʄ/, as examples of "double articulation". These segments have a degree of narrow approximation at the alveolar ridge, where friction is produced, and a degree of wide approximation between the lips, producing no friction. They are, therefore, labialised alveolar fricatives, the correct term used by Fortune (1968, p.24) on his consonant chart, without, apparently, grasping the full theoretical implications of it. "Labialization" "velarization", palatalization", are, as I have already said earlier, secondary modifications of primary articulatory phenomena.⁵⁸

57. This will be explained in the next chapter, where an extended elaboration of terminology to be used in the ensuing phonetic descriptions is systematically set out (pp 92-118)

58. Incidentally, the fricatives /ɬ, ʄ/, and the implosives /ɓ, ɗ/ are here not grouped in a class of their own as the compound consonants so-called, which they were in An Analytical Grammar of Shona. Since Fortune gave a reason then for separating these sounds from the others I rather expected him to explain here why he has changed his classification.

Another point to make is that these fricatives, /ʃ,ʒ/, are not distinguished from /s, z/ merely by the former having lip-rounding and the latter not, as Fortune claims (1968, p.9). They differ also in their places of articulation and in the size of the central passage for the lung air to pass through at the points of articulation. In my own Palatograms (Palatograms (i) - (iv), [s,z] are much farther forward, closer to the teeth ridge than [ʃ,ʒ]; and the latter have a wider air passage than the former. On the acoustic spectrum, [ʃ,ʒ] have a lower energy limit on the frequency scale than [s,z]. All these factors contribute to the perceptual distance between the two pairs of fricatives.

Finally, I think the term "double articulation" should be avoided since it has become well-established in linguistics literature, especially in continental Europe, for describing an entirely different set of phenomena. The French linguist, A. Martinet (1960, pp.22-24, 46-47)⁵⁹, uses "double articulation" to refer to the double segmentation of the continuum of speech into a series of units of grammar or meaning, that is, "morphemes" or, to use Martinet's terminology, "monèmes" - this is the first articulation. Monèmes are themselves capable of being divided into a series of units of sound, that is, phonemes - this is the second articulation. This double articulation accounts for the fact that human language requires only a small number of phonemes to form an infinite number of meaningful utterances. I am suggesting, therefore, that

59. Martinet, A. (1960) Elements of General Linguistics
Translated from French by E. Palmer, Faber, London, 1966.

"double articulation" be reserved for this important characteristic of language, and the term "co-articulation" be used to refer to the articulatory phenomenon represented by the secondary cardinal vowels, for example.

Section 1.6.1 (1968, p.10) is on Individual variation in speech. Fortune writes,

"If we compare men and women, boys and girls, the old and the young, speaking the same language, we find members of these groups and individuals of each group, articulating the phonemes in different ways."

Strictly speaking, however, people do not articulate phonemes, but speech sounds. And the latter are environmentally conditioned, that is, their physical essence reflects the phonetic environment. And although speech sounds differ from person to person, and from occasion to occasion, for intelligibility, the systematic difference between them must be constant.

Phonemes, on the other hand, are abstract constructs intended by the linguist to enable him to handle the multitudes of actual sounds occurring in speech. As Fortune later says in the same section, the linguist is concerned with describing these systematic differences. Speech sounds must be "placed" in relation to other sounds in the same system. Edward Sapir (1925)⁶⁰ puts it in this way,

"A sound that is not unconsciously felt as 'placed' with reference to other sounds is no more a true element of speech than a lifting of the foot is a dance step unless it can be 'placed' with reference to other movements that help to define the dance."

60. Sapir, E. (1925) "Sound Patterns in Language" Language, 1, pp.37-51. Reproduced in Makkai (ed.) op.cit., pp.13-21

1.4.2.1. Fortune's Phonemic analysis of the stops.

The oral stops are discussed in section 1.6.3. His analysis and classification of these phonemes calls for comment. For Fortune, (1968, pp.12-24), Central Shona has three "depressor" (for me "breathy-voiced") stop phonemes, namely:

/b̥, d̥, g̥/.

Their distribution is as follows:

/b̥/ has the following allophones:-

- (i) [b̥], the single syllable onset;
- (ii) [b̥], the second element in the cluster [mb̥];
- (iii) [b̥], the first element in the cluster [b̥y].

/d̥/ has the following allophones:-

- (i) [d̥], the single syllable onset;
- (ii) [d̥], the first element in the affricates [d̥z, dz̥, dz̥];
- (iii) [d̥], the second element in the cluster [nd̥].

/g̥/ has the following allophones:-

- (i) [g̥], the single syllable onset;
- (ii) [g̥], the second element in the cluster [ŋg̥].

This covers the breathy-voiced set. Fortune also recognises three plain-voiced stop phonemes:

/b, d, g/, with the following distribution:

- (i) [ɓ, ɗ, ɠ], i.e. the implosives, occurring as single syllable onsets;
- (ii) [b, d, g], occurring as the second element respectively in [mb, nd, ŋg].

Fortune's interpretation gives us three breathy-voiced phonemes, and three plain-voiced ones, a total of six.

However, it is this last claim that I cannot

agree with: that [ɖ], an implosive, and [d], an explosive, are allophones of the same phoneme /d/ - which, strangely enough, never occurs as a single syllable onset. The same interpretation is also extended to the bilabial and velar plain-voiced stops.

Fortune agrees with Bloch and Trager (1942,p.42),⁶¹ among others, who insist on the criterion of phonetic similarity being met before two sounds can be treated as allophones of the same phoneme. To quote Bloch and Trager,

"If two or more sounds are so distributed among the forms of a language that none of them ever occurs in exactly the same position as any of the others, and if all the sounds in question are phonetically similar in the sense of sharing a feature of articulation absent from all other sounds, then they are to be classified together as allophones of the same phoneme," (my own emphasis).

As it happens, the explosive stops of Central Shona, both breathy- and plain-voiced, share more phonetic and phonotactic characteristics with each other than either of them with the homorganic implosive stop. It is because I think Fortune, in classifying the plain-voiced implosive and explosive in the same phoneme is, among other things, violating the criterion of phonetic similarity as defined by the American structuralists, that I am proposing an alternative phonemic interpretation.

1.4.2.2. An Alternative Phonemic Analysis of the Stops.⁶²

I think the following analysis meets the condition

61. Bloch, B. and Trager, G.L. (1942), op.cit.

62. This interpretation is elaborated in the section on the stop phonemes in chapter 5, (pp. 235-240).

of phonetic similarity and is more intuitively satisfactory than Fortune's. I accept that there are three breathy-voiced phonemes in Central Shona:

/b̤, d̤, g̤/, with the following distributional allophones:-

- (i) [b̤, d̤, g̤], functioning singly as syllable onsets;
- (ii) [mb̤, nd̤, ŋg̤], i.e. combining with the homorganic nasals in the so-called naso-oral complexes with breathy-voice;
- (iii) [mb, nd, ng], i.e. in combination with the homorganic, plain-voiced nasals in naso-oral complexes;
- (iv) [b̤v, d̤z, d̤ʒ, d̤ʒ], i.e. in affricative clusters with homorganic or semi-homorganic fricatives.

Finally, this interpretation recognises two implosive phonemes:

/ɓ, ɗ/, previously treated, by Fortune, as allophones of the plain-voiced "phonemes"

/b, d/.

These two implosive phonemes do not occur in clusters with any other consonants but act as single syllable onsets before all vowels. There is no velar implosive in Central Shona, contrary to Fortune. Altogether, this alternative interpretation gives us three breathy-voiced explosive phonemes, and two plain-voiced implosive phonemes. Together with the three undisputed voiceless phonemes, this gives us eight plosive phonemes as opposed to Fortune's nine. The second interpretation is therefore more economical.

In chapter 5 I shall argue that besides the fact that /ɓ/ and /b̤/, and /ɗ/ and /d̤/ are in contrastive distribution:

- (a) /b̥ḁ/, ideophone of suddenly catching sight of
vs
/b̥ḁ/, ideophone of finishing food or of hitting
something.
- (b) /-ɖḁɖḁ/, "be insolent",
vs
/ɖḁɖḁ/, "duch" (n.5).

and should therefore be two independent phonemes, implosives and explosives are so phonetically dissimilar that they should not be classified in the same phoneme. Their dissimilarity consists in the speech mechanism employing different airstreams - ingressive and egressive - to produce the two sound types. As I have already noted above, after all the other criteria have been satisfied, in English [ŋ] and [h] are assigned to different phonemes because they are phonetically dissimilar. This criterion is the acid test.

In 1.6.6., where Fortune describes the glottal fricative /h̥/, he says something which could have helped him in his analysis and classification of the oral stops.

"/h̥/ should clearly not be classified with the voiced depressor spirants. It is very different from them in its mode of production and it does not enter into clusters as they do," (1968, p.16).

This is a good argument for not classifying implosives in the same phonemes with explosives and also for separating nasals from plosives - although I do not agree that /h̥/ should be isolated from other fricatives.

In 1.6.7 (1968, p.18) Fortune writes,

"The articulation of /v/ differs from that of /v̥/ ... /v̥/ is a spirant and its point of articulation is the meeting of the inner surface of the lower lip and the upper teeth,"

that is, it is also labio-dental, although he says they differ in "place and manner of articulation". I agree that they differ in manner of articulation, because one is a fricative while the other is a frictionless continuant. In my pronunciation both are labio-dental articulations. They differ in their duration too. Whereas /y/ can be prolonged indefinitely, /v/, which I shall symbolise as /ʋ/, has a duration similar to that of the semi-vowels, very short. The second part of the chapter on semi-vowels is devoted to /ʋ/ because of these articulatory and consequently, acoustic similarities between these segments.

Section 1.6.10.1 (1968, pp.20-21) is on "Cw clusters by which Fortune means combinations of other consonants with the semi-vowel /w/.

"The phoneme /w/ has a large number of allophones due to the fact that /w/ is assimilated to the preceding consonant. The allophone is determined by the preceding consonant. If the preceding consonant is a labial, the allophone of /w/ will be a velar consonant; if alveolar or prepalatal, the allophone will be a velar consonant +[w]..."

This is very confusing. In Fortune (1955, p26) assimilation is defined as a process whereby a sound becomes similar to a neighbouring sound. Fortune does not give a different definition in the present book. So it is strange that he should regard a combination of a labial and a velar, for example:

/px, bʏ/, in /pxére/ "children" and /bʏe/ "stone" as evidence that /w/ has become assimilated to /p/ and /b/, or that a sequence of an alveolar and a velar +/w/ is also an instance of the assimilation of /w/ to the

alveolar, for example:

txw, in /txw^áana/, "little children" (Zezuru).

We are also told that "the extent of the raising of the back of the tongue varies dialectally"! The so-called allophones of /w/ after labials are listed as [k, g, ŋ, x, ɣ].

"In this interpretation we find the following as allophones of /w/: [kwx, g~ɣ, ɣ, Kw~xw, gw~ɣw, ŋw, ɣw, w]... This means that features which differentiate phonemes elsewhere cease to do so in this environment... here features of allophones are determined by environment and are not in contrast." *(own emphasis A.P.)*

But allophones are always environmentally conditioned.

I do not see why Fortune speaks of this particular environment like that.

The second point to make is related to his claim that features which differentiate phonemes elsewhere in the system cease to do so here. Fortune gives a number of examples to support this rather strange claim, unwittingly including a pair of contrasting examples:

/itʃk^á/, "fear!"

/idzɡ^á/, "eat!"

In other words these two words are distinguished from each other by what Fortune has just listed as allophones of /w/, namely [k] and [g]. But as we know, allophones of one phoneme do not occur in the same phonetic environment. There are more counter-examples which go to disprove that certain phonemes, as it were, suspend their contrastiveness here:

(i) /-dʒɪp^á/, "be choked" vs. /dʒɪb^á/, "mucus" or "a cold"

(ii) /-rɛʃx^á/, as in [p^árɛʃx^á], "an impropriety has been committed".

vs

/-rè^ωz̥^ωá/ as in [á^ωrè^ωz̥^ωá], "it (the baby) has been comforted".

((ii) as in the Gutu Karanga area).

(iii) /-k^ωírá/, "climb, ride" vs. /-x^ωírá/, "pound in..." or "for ...".

The last example is particularly damaging to Fortune's analysis: [k^ω] and [x^ω] are both velar and voiceless, and are adjacent to /i/, i.e. have the same environment.

How can this happen if they are allophones of the same phoneme /w/? And what is the conditioning factor by which we get [k] in one instance and [x] in another?

This interpretation simply will not do.

Fortune concludes this section by conceding,

"The fact that there is so much assimilation (I have argued that there is no assimilation) gradation and free variation (I have demonstrated that there is contrast) between the realizations of /w/ is an argument for the interpretation offered here" (Fortune, 1968, pp.22-23).

There is really no argument. Then, "The treatment of clusters, of which five types have been given, is frankly inconsistent," which is an understatement.

By the time I come to discussing the so-called /Cw/ clusters at the end of chapter 5, I shall have established that all the segments listed by Fortune as allophones of /w/ are phonemes in their own right, that is /k, x, g, ɣ, ŋ, ɲ/. In that section I shall argue that the complexes which are symbolised as /kw, g^w, ɣ^w, ŋ^w, mw etc/ in Doke's and Fortune's phonetics and phonology, and treated by them as sequences of /C/ and /w/, are not sequences at all. This is how they are symbolised in the ordinary orthography, and this is what they are

describing - the orthography. No Shona speaker pronounces these complexes as sequences. For example, in [-k^wirá], "climb", the velar closure for [k], and lip-rounding, occur simultaneously and, auditorily, one does not hear [k-w], i.e. a sequence, but [k^w], i.e., a labialised voiceless velar stop. I shall, therefore, from now on, treat all so-called clusters of /Cw/ as instances of the labialization of the /C/ in question. Therefore I shall recognise a set of phonemes to be symbolised /C^w/, not /Cw/.

My interpretation, which at first was based on my kinaesthetic and auditory impressions, has the backing of acoustic evidence. The speech segmentation does not have a separate portion for the /w/; spectrograms and mingo-grams have no separate tracings for /w/.

Finally, all combinations of C and so-called allophones of /w/: / px, bɣ, pk, tx, dʒg, etc/ will be treated as sequences of independent phonemes, which they are auditorily, kinaesthetically.

1.4.2.3. Fortune's analysis of Suprasegmental phenomena.

The phonemic analysis of the segmentals just discussed was also extended to the suprasegmental phenomena of tone and length. Fortune's analysis of tone and length is decidedly better than that of the segmental phonemes.

"Pitch and duration are used in Shona phonology to provide contrasts. The phonemes of length and tone are suprasegmental because they are carried by the syllable composed of segmental phonemes (so-called because they follow one another in regular patterns of alternation and can be distinguished from one another in their sequences," (1968, p.27).

Fortune's definition of the domain of the supra-segmental phoneme is faultless within the framework of phonemic analysis. Generative phonologists, as we shall see later, assign tone to the [+syllabic] segment rather than to the syllable as a whole.

What I would like to question is what Fortune says about the serial order of segmental phonemes, in parenthesis above. Segmental phonemes follow one another in the manner described by Fortune only in transcription and orthography. In actual speech there is a continuity. Gimson (1962, p.290)⁶³ speaks of the "mutual influence which contiguous elements have on each other", while Dwight Bolinger (1968, p.43)⁶⁴ declares that

"phonemes are affected by the company they keep, that is, they do not keep their identities like a row of bullets laid end to end."

Edward Sapir (1933)⁶⁵ refers to the "intricate mathematics of significant relationships" between sounds in different phonetic contexts. One only needs to listen to people speaking or to look at acoustic records of speech on spectrograms and mingograms to appreciate what these authors are saying and what Fortune has missed: that speech is

63. Gimson, A.C. 1962, op.cit. "If, therefore, the utterance is analysed in terms of a sequence of phonemes, account must be taken of the phonetic continuity and merging of qualities by describing the mutual influence which contiguous elements exert upon each other- in other words, tendencies towards 'assimilation' have to be noted."

64. Bolinger, D. 1968, Aspects of Language, Harcourt, Brace and World Inc.

65. Sapir, E. 1933, "Language" In Encyclopaedia of the Social Sciences, Vol.9, pp.155-169, New York, Macmillan Co. Also in D.G. Mandelbaum (ed.) Edward Sapir: Culture Language and Personality: Selected Essays, pp.1-44. University of California Press, Berkeley and Los Angeles 1970.

on-going.

In section 1.7.2 Fortune discusses Phonemes of tone. The phonemes of tone "cannot be identified save in contrast. This contrast is supplied normally by the succession of high and low tones:

/nèsádzà/ 'with porridge'

/némvurà/ 'with water'.

The identity of every tone is maintained by that tone being kept distinct from other tones in the neighbourhood in words. But to establish the systematic tones in a language under investigation using contrast we need to use minimal pairs. This works very well for Central Shona. In such pairs the tones enter into paradigmatic relationships:

/gùrú/ "huge" (adj. cl.5)

vs

/gùrú/ "3rd stomach of ruminant".

The meaning difference between the two words can be related to the tone of the initial syllable of each, i.e. [+HIGH] for the first and [-HIGH] for the second. These are the two systematic tonal phonemes of the language. In the examples given by Fortune, i.e. /nèsádzà/ and /némvurà/, the tones are in syntagmatic relationships. What we can see there are the phonotactic possibilities of these tones. This is not how tones are discovered: the difference between one tone and the next may be the result of conditioning by different consonants. So how can /nèsádzà/ help us to discover that Shona has only two contrastive tones?

In section 1.7.2.2. (1968, p.30) Fortune discusses tonal Assimilation and Dissimilation.

In this⁶⁶ review I shall comment only on one of the points made by Fortune. "Assimilation occurs in sequences of substantives in substantive phrases (save that demonstratives are not affected)..." He also adds that "Assimilation does not occur after a pause." And to illustrate this last point he gives the sentence (which I transcribe using I.P.A. symbols): [bàbá wàngù mùdíkì ≠ ákàḡá nāmà zínḡí ≠ nókumḡá hāḡá úzínḡí ≠ pàtṣìḡàrò ítṣì]⁶⁷ (Fortune uses ≠ to indicate a pause). Then he says "In this example [nāmà] is assimilated as [nāmà], but [pàtṣìḡàrò ítṣì] is not changed," (1968, p.30).

First of all, it is true both that [pàtṣìḡàrò ítṣì] "on this stool", is not assimilated and that it comes after a pause. But Fortune has already said that assimilation does not occur in demonstratives - "save that demonstratives are not affected" - and [pàtṣìḡàrò ítṣì] is a demonstrative phrase. This fact makes this particular phrase an unlikely choice for illustrating the claim that tonal assimilation does not occur after a pause.

In conclusion to this section, Fortune gives a number of phrases exemplifying assimilation and dissimilation:

/wàrúmé wá rē: f ú /, "tall men"

/wàrúmé wá rē: f ú /, "the men are tall"

(the second word has underlying tone pattern /wàrēfú/).

I must disagree with his rendition of the second example.

The phrase as it stands could not be a Shona utterance.

To convey the English meaning "the men are tall" the

66. Tone is discussed in chapter 7 (pp 385-453) below, where I have made some comments on Doke's and Fortune's handling of tone in Shona.

67. "My uncle ate a lot of meat and drank a lot beer (sitting on this stool."

phrase must have the following form:

/v̥arúmé vátʃò v̥arèfú /, that is, to make the phrase definite, "the men", one must introduce a particle.

In section 1.7.3 Fortune discusses Phonemes of Length. There being no minimal pair of Shona words distinguished by corresponding syllables having different degrees of length, I find "phonemes of length" strange. There is, indeed, length on penultimate syllables in words, especially at the end of a phrase or clause, and Fortune gives a number of examples (1968, p.31). But nowhere does he give examples of words distinguished by syllable length, as

/gúru/ "large" (adj. 5)

and

/gùru/ "3rd stomach of ruminant" (n.5) are distinguished by the tonemes /+high/ and /-high/ on the initial syllables, or, as

/-tamba/, "play" (v.1.)

and

/Kamba/, "tortoise" (n.9/10) are distinguished by /t/ and /k/, their respective initial segmental phonemes.

In section 1.7.3.2 Fortune discusses Length patterns in words. He gives the example

/æ:ndà/, as an example of penultimate syllable lengthening, rendering it as English

"he went", that is, the past tense. This, however, is not the correct rendition. The tone pattern would have to be different, and the segmental structure

of the form altered:

/ákáé:ndà/, "he went". As it stands,
Fortune's example is a relative clause:

"he who has gone".

I am aware of the fact that Fortune is here concerned mainly with the suprasegmental length. However, I think he is creating problems for the grammar in his phonology by giving incorrect renditions to certain constructions in this section.

Conclusion

The preceding remarks on Fortune's phonetics and phonology - and there are more such remarks in a later chapter - must be seen in their proper perspective. Fortune is not a phonetician. And I believe he would like to be remembered as a syntactician and semanticist. This raises the question why he finds it necessary to preface his grammars with chapters on phonetics and phonology - always referring his readers to C.M. Doke (1931)⁶⁸ for more detailed discussion.

I think part of the reason is that Fortune knows the importance of an accurate phonetic analysis of a tone language such as Shona before a tolerable grammar can be written. Henry Sweet (1891)⁶⁹ wrote "It is now generally recognised, except in hopelessly

68. Doke, C.M. (1931), op.cit.

69. Sweet, Henry (1891) A New English Grammar, Excerpts in E.J.A. Henderson (ed.) The Indispensable Foundation: A selection of the writings of Henry Sweet, O.U.P. 1971.

obscurantist circles that phonology is the indispensable foundation of all linguistic study..." And Edward Sapir (1929)⁷⁰ echoed these views when he referred to phonetics as "... a necessary prerequisite for all exact work in linguistics.."

This means that the criticisms that I have made of Doke's and Fortune's phonetic studies of Shona have serious implications for the grammar which such studies are intended to support. The last criticism above, about Fortune's incorrect rendering of some Shona tense formatives, is a case in point. If these formatives are wrongly rendered in the phonology, and then treated differently in the grammar, the resulting grammar will lack cohesion. It is true that all grammars leak, but they leak for different reasons; Fortune's grammar probably because of inadequate phonetics. Doke explained the limitations of his study as having resulted from (a) a shortage of time to conduct extensive enough research; (b) limited financial resources; (c) a lack of sophisticated research equipment. That was in 1929. And Fortune has been quoting Doke (1931) as the definitive work on Shona phonetics since the mid-50s. There has been no alternative. However, credit will always be due to Doke, Fortune and Hannan⁷¹ the lexicographer, for introducing Shona linguistic studies to the international scene during a period of intense

70. Sapir, E. (1929) "The Status of Linguistics as a science", in D.G. Mandelbaum (ed.), op.cit., pp.65-77.

71. Hannan, M., S.J. (1974) Standard Shona Dictionary, Rhodesia Literature Bureau.

and extensive debate and research.

The ensuing study of the phonetics and part of the phonology of the Karanga will, I hope, tackle some of the problems faced by Doke and Fortune and serve as an example of the way a more extensive study of all the dialects using modern laboratory and field techniques and theories can be conducted.

CHAPTER 2

2.0 PHONETIC ANALYSIS AND CLASSIFICATION: A THEORETICAL FRAMEWORK

Before attempting a detailed description and classification of the distinctive sounds of the Karanga dialect, it is necessary to give an outline of what seems to be a reasonable and rigorous enough general phonetic framework to serve as a guide in the exercise. This outline is placed here so that when, later on, certain terms are used and judgments made, the reader will have already familiarised himself with the theoretical¹ basis from which they derive.

2.1 On the "Speech Chain"

In order to facilitate description, the speech chain has been divided into a number of phases:

- (i) the physiological phase:- (a) the initiatory stage;
(b) the articulatory stage;
(c) the auditory stage.
- (ii) the acoustic phase.

There is another phase, called the "linguistic level" by Denes and Pinson (1963)². But I have left that out because it concerns the mental processes which precede, indeed underlie, speech. These activities,

-
1. I have benefited from the lectures on "Phonetic Analysis and Classification" given in the period 1972-1974 by members of staff in the Department of Phonetics and Linguistics (Phonetics Section), University College, London.
 2. Denes, P.B. and Pinson, E.N. (1963) The Speech Chain, Bell Laboratories.

important though they are, lie outside the scope of this thesis.

'Physiology' refers to the study of what the muscles and bones of the body do, as opposed to their structure, which falls in the domain of 'anatomy'. In the discussion of the articulatory stage of speech production, the emphasis will be on what the muscles and bone structures of the speech tract do, their function in the production of sound waves.

2.2 The Physiological Phase

Initiation.

One of the functions of the cavities involved in speech-production is to provide an air-stream, that is, initiate an airflow. Following J.C. Catford (1970)³; I shall discuss this under two sub-headings, namely the location and the direction of the initiatory activity. Under direction there are two contrastive possibilities: egressive and ingressive. An egressive air-stream flows from the vocal tract, whereas an ingressive air-stream flows into the vocal tract.

There are three major locations of initiatory activity: the lungs, the larynx and the oral cavity. Initiation that begins in the lungs is called 'pulmonic', while that in the larynx is called 'glottalic', (some call it 'pharyngal'). Initiation in the mouth takes place at the velum and is consequently called 'velaric'.

3. Catford, J.C. (1970) "The articulatory possibilities of man" In Malmberg (ed.) Manual of Phonetics, North-Holland.

2.2.1 The major Initiation Types

Direction and Location

- (a) Pulmonic egressive. The volume of the lungs is decreased, and the air in the lungs is pushed outwards through the trachea, the larynx, the pharynx and out through the mouth or nose or both. This initiation type is used in the majority of the languages of the world, and for most of them it is the only one used.
- (b) Pulmonic ingressive. This initiatory type is the opposite of pulmonic egressive. The volume of the lungs is increased, thereby producing rarefaction of the enclosed air. This results in a potential or actual flow of air into the vocal tract through the mouth or nose or both.
- (c) Glottalic egressive. In this initiation type the glottis is closed and a stricture of some degree is made in one of the supra-glottal cavities. The larynx moves upwards, generating positive air pressure in the pharynx and mouth, and a potential outward flow of air. Glottalic egressive initiation is reported in the Ndau dialect of Shona by Mkanganwi (1973)⁴. The sounds produced using this air-stream are generally known as 'ejectives'. Catford (1939)⁵ calls them 'glottalic pressure stops'.

4. Mkanganwi, K.G. (1973) An Outline of the Morphology of Substantives in Ndau, with a preliminary note on Ndau Phonology. Unpublished M.Phil. thesis, University of London.

5. Catford, J.C. (1939) "On the Classification of Stop Consonants" In Jones and Laver (eds.) op.cit.

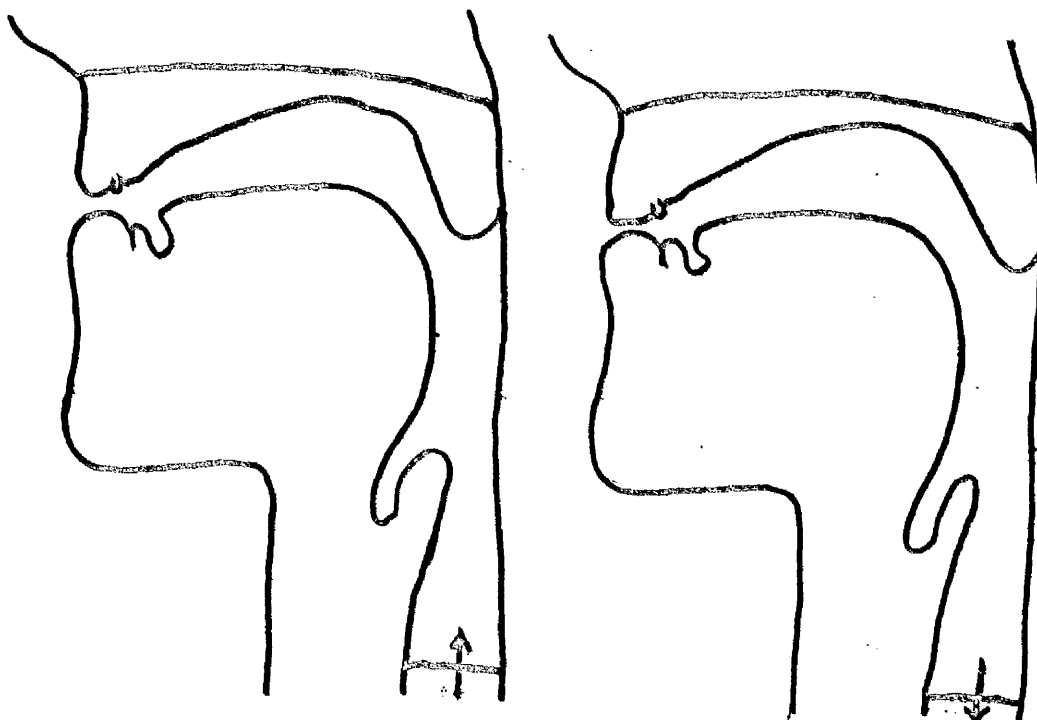


Diagram 1. showing (a) Glottalic egressive initiation. (b) Glottalic ingressive initiation.

(d) Glottalic ingressive. The vocal folds are held loosely together and the larynx as a whole moves downwards, generating negative air pressure, that is, rarefaction in the pharynx and mouth, and a potential or actual inward flow of air. Shona has sounds produced by this mechanism. They are usually called "implosives", Beach (1938)⁶ called them "Kaffir [b, d, g]," while Catford (1939) called them "glottalic suction stops". Both auditorily and articulatorily, 'implosive' seems to be the most satisfactory descriptive term.

(e) Velaric egressive. The initiatory stricture here is made by the back of the tongue coming into firm contact with the velum. In addition to this stricture,

6. Beach, D.M. (1938) The Phonetics of the Hottentot Language, Cambridge; Heffer.

there is an articulatory stricture farther forward in the mouth. A forward and upward movement of the central part of the tongue produces positive air pressure behind the second stricture. This is the spitting mechanism, so-called. It is also sometimes known as the reverse click.

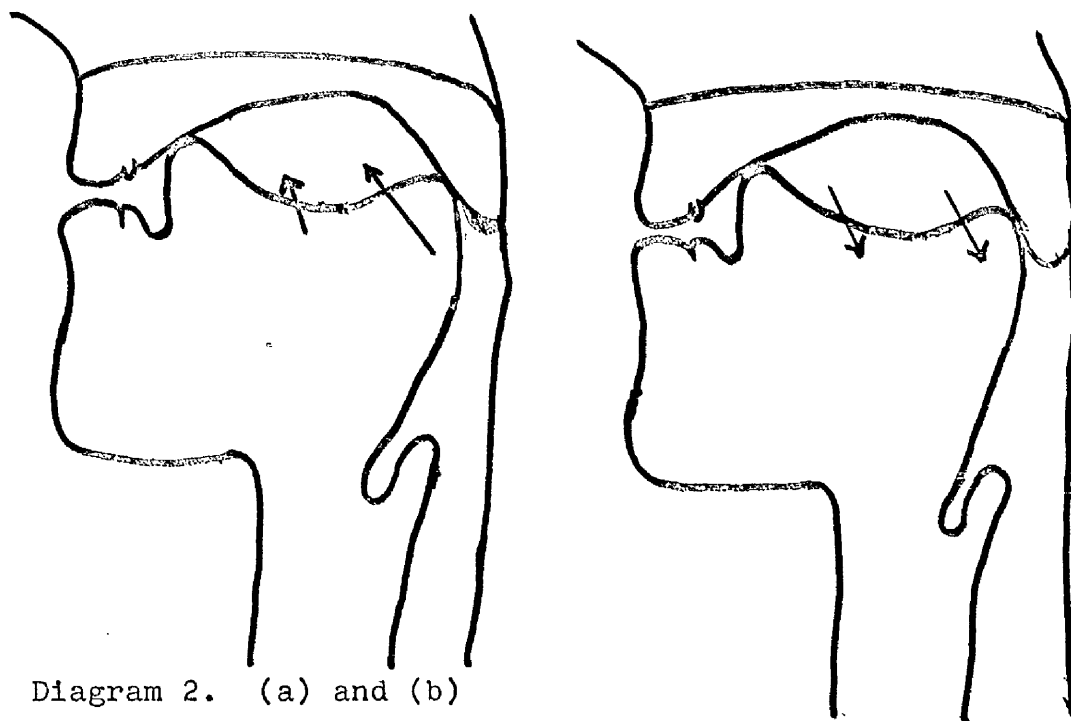


Diagram 2. (a) and (b)
showing (a) the velaric
egressive air-stream

and (b) the velaric ingressive
air-stream mechanism

(f) Velaric ingressive. As in the velaric egressive airstream above, there is a firm closure between the back of the tongue and the velum, and a second stricture farther forward in the mouth. There is, however, a downward and backward movement of the central part of the tongue, which creates negative pressure, i.e. rarefaction, behind the articulatory stricture. The speech sounds so produced are called 'clicks', widely used in the Nguni group of Bantu, and randomly in the Selukwe, Gwelo and Que Que varieties of the Karanga dialect:

/mùʔirò/, "a whip" (n.3)

/-ʔupà/, "rinse mouth"

/maʔimbi/, "edible caterpillars", using the alveolar or dental click.

2.3 The articulatory stage.

There are four cavities which play an important role in the production of speech sounds. As Diagram 3 shows, these are:

1. the oral cavity;
2. the nasal cavity;
3. the pharyngeal cavity; and
4. the pulmonic cavity.

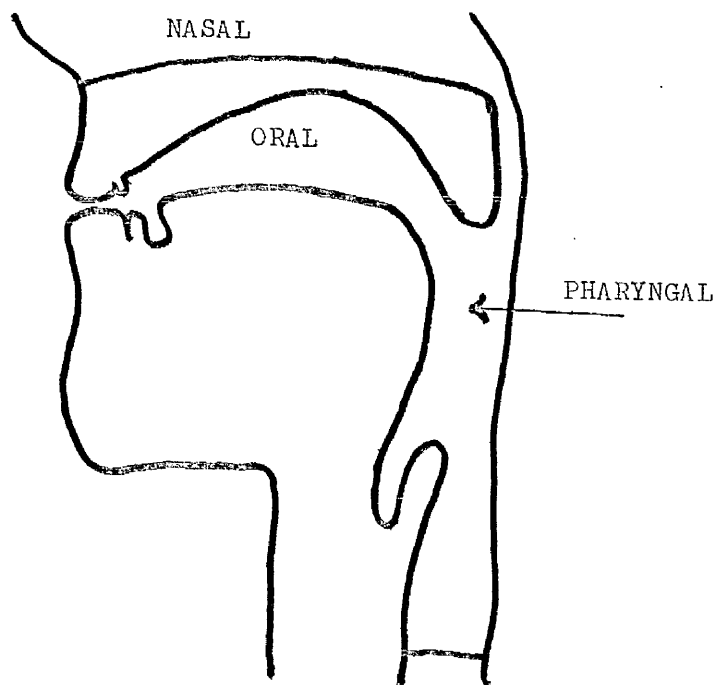


Diagram 3. The cavities used in speech production.

These cavities are divided into two general

categories: (a) the sub-glottal, and (b) the supra-glottal cavities. With the exception of the pulmonic cavity, which is sub-glottal, all the other cavities are supra-glottal.

(a) Functions of the cavities. The general functions of these cavities are of two kinds: (i) to provide an air-stream. The pulmonic cavity provides the air-stream used in the production of the majority of the sounds of speech. The range of sounds produced using the air-stream provided by the other cavities is by far smaller. The nasal cavity is a rigid structure and is therefore unable to act as an initiator.

(ii) To produce sounds: The size and shape of these cavities can be modified, producing many different sound qualities. This is an exclusive function of the supra-glottal cavities, particularly the oral cavity, which houses these very flexible organs: the tongue, lips, and the uvula. Although the larynx acts as an initiator of the glottalic egressive and ingressive air-streams, its primary function is phonatory in conjunction with the pulmonic and glottalic airstreams, both egressive and ingressive. In Shona there are three distinct possibilities regarding the phonatory function of the larynx: it can be wide open, for voiceless sounds; the vocal folds may be held loosely together throughout their entire length for plain-voice; or, the arytenoid cartilages may be held apart, while the ligamental vocal cords are vibrating ,

for breathy-voice or murmur.⁷

Although the nasal cavity acts primarily as a resonator, there is likely to be some nasal friction for voiceless or devoiced nasals: [m̥, n̥].

In the description of speech sounds, it is necessary to arrange these cavities in some order of importance using phonetic criteria. There are some sounds which utilise more than one cavity simultaneously. For example, which descriptive label should be given to the articulation represented by the symbol [ã]? Is it an oralised nasal or nasalised oral? If we give it the latter label, we are implying that the oral cavity plays a more decisive role in the articulation of the vowel than the nasal; and conversely, if we choose the former label. In either case, our choice of label must of necessity have phonetic plausibility: it must be based on an accurate observation of the activities involved in the production of the sound in question.

2.3.1 The condition of the cavity.

We can try to resolve the problem posed above by referring to what the air-stream does in the cavities involved in the articulation of the sound in question. We use three terms, all referring to the condition of the cavity. The cavity can be in any one of the following conditions:

(a) Active: this refers to a cavity into which the air

7. Ladefoged, P. (1971), op.cit., pp.6,8.

used in an articulation enters and passes through;

- (b) Semi-active: this refers to a cavity into which air enters but does not pass through during the articulation of a sound;
- (c) Inactive: an inactive cavity is one into which air does not enter during the pronunciation of a sound.

The oral cavity is active during the pronunciation of all oral and nasalised oral sounds; semi-active during the pronunciation of all nasal sounds except /N/, the uvular nasal - during which it is inactive; and inactive during the hold stage of the glottal stop. The nasal cavity is always either active or inactive: active during all nasal or nasalised oral sounds; and inactive during the pronunciation of all non-nasalised oral sounds. The pharyngeal cavity is almost always active. It is semi-active during the hold stage of the two uvular stops [q, G], and inactive during the hold stage of the glottal stop [ʔ]

These distinctions are useful and important because they can be used as well to determine the relative importance of these cavities for purposes of describing an articulation, as for dividing speech sounds generally into the two broad categories:

- (a) continuant
- (b) non-continuant. During the pronunciation of continuants either the oral cavity or the nasal cavity or both are active, whereas during non-continuant the nasal cavity is inactive and the oral, semi-active for a period of time.

During the pronunciation of any speech sound, a cavity will rank high, depending on whether it is active, semi-active or inactive, in that order. That is to say, in describing a sound we shall use the term 'oral' if that cavity is active while the nasal is inactive. And the label 'nasal' will be used if the nasal cavity is active while the oral is semi-active or inactive during the pronunciation of a particular nasal sound.

So one of the criteria that will be used for arranging these cavities in an order of importance for descriptive purposes is that of condition.

2.3.2 The Degree of stricture

The second criterion used may be termed degree of stricture. During the pronunciation of a sound, the organs of speech may assume any one of three "postures":

- (a) complete closure, that is, there is at some point in the speech tract firm contact between two organs of speech, cutting off the air-stream;
- (b) narrow approximation, that is, two organs of speech come very close to each other, but without actually making contact, thus allowing air to go through fricatively;
- (c) wide approximation, that is, the organs of speech are wide apart, and air flows unimpeded from the lungs, there being no ingressive continuants.

The two degrees of approximation (b) and (c) are established by auditory means. The stricture of narrow

102
approximation ranks higher than that of wide approximation because it results in audible friction, whereas some of the articulations involving wide approximation are inaudible. With the exception of lip closure during quiet breathing, closure between organs of speech ranks as the highest stricture because it represents the most radical departure from their position of rest.

The strictures of narrow and wide approximation apply to the category continuant, fricative, and frictionless continuant respectively; whereas complete closure applies to the category non-continuant.

The criteria adduced so far for determining the relative importance of the four cavities do not help us to resolve the problem posed above, namely, whether to characterise the articulation [ã] as an oralised nasal or nasalised oral. This is so because during the articulation of [ã] both the oral cavity and the nasal cavity are active, and they have the same degree of approximation, that is, wide approximation.

2.3.3. Articulatory Potential

When two cavities both playing a part in the production of a sound have the same degrees of stricture and are in the same condition, we rank one of them as being higher than the other because of the criterion known as articulatory potential, our third criterion. We decide this by asking the question, which one of the two competing cavities is capable of producing a wider range of speech sounds than the other? In this case the oral cavity

will have to be ranked higher than the nasal, so that [ã] will now be characterised as a nasalised oral.

2.4 Modes of egress.

Having described both the conditions of the cavities involved in speech production and the strictures which the organs of speech assume in the various cavities, what remains now is to specify the ways in which the air-stream leaves the vocal tract. The phrase "mode of egress" refers to the direction and path followed by the outgoing air involved in the production of speech sounds. Thus the vowel sound [e] has an oral mode of egress, whereas the consonant [n] has a nasal mode of egress.

Air that escapes through the oral cavity has two possibilities: (a) median or lateral egress, and (b) if median, slit or grooved egress. The air can issue via the midline of the oral cavity, with or without friction. For example, Shona [ɖ] has a median fricative mode of egress; whereas [a] also has a median, but frictionless mode of egress. Similarly, air can also issue at the sides of the mouth, i.e. laterally. For some languages lateral egress may be either frictionless or fricative, e.g. Ndebele⁸ /l/ in /úkúlàmbà/, "to starve", /ɽ/ in /úkúṭàṇà/, "to be mad". But laterality in Karanga has only the fricative possibility:

/ɽ/ in /Kùṭá/, "to fear"

/k/ in /Kùḡá/ "to eat"

8. Bantu language of Rhodesia, belonging to the Nguni cluster.

104

The difference between median and lateral modes of egress arises from the absence of median obstruction of the oral cavity in the former and its presence in the latter.

The slit-versus-grooved dichotomy is applicable only to the stricture of narrow approximation. In a slit mode of egress, the tongue is flat across from side to side, for example in English [θ]. In a grooved mode of egress, the tongue is furrowed along the centre, as in the Shona fricative [s].⁹ This difference in the shape of the tongue accounts for the stridency of [s], which is missing in [θ]. If we assume that there is equal air pressure during the pronunciation of the one sound as during that of the other, a grooved stricture will result in a more strident sound than will a slit stricture.

2.5 (a) Minor Articulatory Features

These articulations take place in conjunction with major articulatory features. Since the two types occur simultaneously, the points of articulation are at different locations in the speech tract. Major and minor articulatory features are of two kinds. First, we can have the two strictures in the same cavity, one ranking higher than the other in terms of the criteria established above.

9. See the first Palatogram in Vol. II.

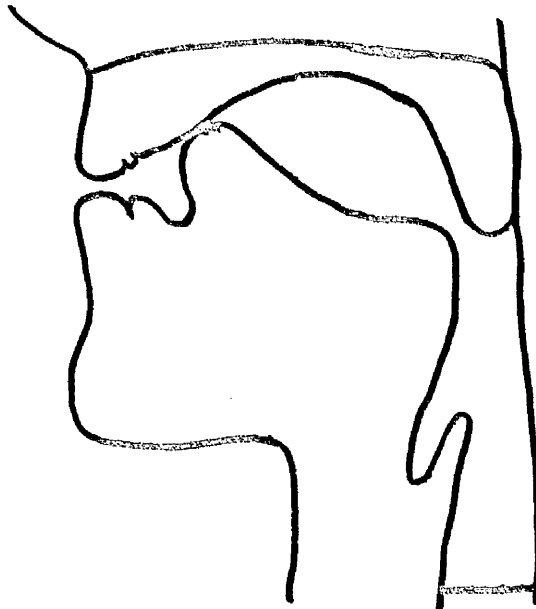


Diagram 4. Articulatory 'postures' for [ɬ].

For example, both strictures for [ɬ] are in the oral cavity, as shown in Diagram 4. The articulation is called a velarised alveolar lateral because the stricture at the alveolar ridge, a complete closure, ranks higher than that between the back of the tongue and the soft palate, a stricture of wide approximation.

Secondly, we may have the same degree of stricture in two cavities which are also in the same condition, that is, active, semi-active or inactive. But one of the cavities will rank higher by the criterion of articulatory potential. The articulation [ɬ̥], discussed above, is a case in point.

There is a problem with regard to [tʰ], often called the 'glottalised' alveolar ejective, a characterization which J.C. Catford (1970)¹⁰ objects to. This is because he regards the stricture at the glottis as articulatory. Calling [tʰ] - see Diagram 5 - 'glottalised'

10. Catford, J.C. (1970) op.cit., p.314 in Malmberg (ed.)

obscures this distinction for Catford, and I agree with him, because the activity at the glottis is thereby being identified in kind with that of the front of the tongue in sounds which are correctly described as 'palatalised', or that of the soft palate in sounds described as 'nasalised', for example. I agree that "glottalised" is gratuitous and misleading in this case, but I have not come across this usage; nor does Catford say where he has encountered it.

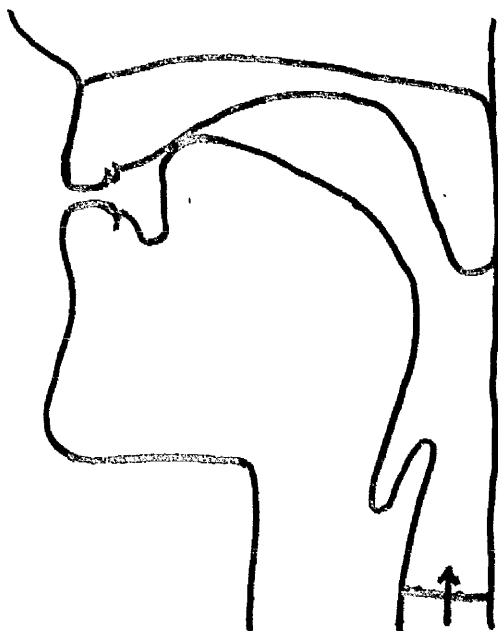


Diagram 5. The ejective [tʰ].

(b) Coarticulations. These are articulations of equal value and they are different from the major and minor combinations just discussed. The articulations involve two equal degrees of stricture in the oral cavity. This is the case in the pronunciation of the first secondary cardinal vowel [y], where we have two equal degrees of stricture: one at the lips, which are rounded; i.e. have a degree of wide approximation; and the other between the palate and the raised front of the tongue,

also a stricture of wide approximation.

2.6 The Auditory Stage.

Auditory phenomena are produced by the various ways in which the organs of speech modify the mode of flow of air into, and out of, the speech tract. There are major and minor auditory phenomena.

2.6.1. Major auditory phenomena.

(i) Plosion. Plosion is a noise that results from the balancing up of two unequal pressures of two bodies of air separated by a closure. There are two types of plosion: explosion and implosion. In the former case, air is released and expelled out of the speech tract (mainly) through the mouth. For implosion, there is an outer closure (in the case of Karanga, the closure is either bilabial or alveolar) and the vocal cords are positioned for producing voice. A sudden depression of the larynx, by enlarging the supraglottal cavities, rarefies the imprisoned air, so that a hollow sound is heard when the outer closure is released.

Plosion is a feature of the category of sounds designated by the term "non-continuant."

(ii) Friction. Friction results from air flowing at a sufficient rate through a narrow passage. Its ingredients are, first, a stricture of narrow approximation; and, secondly, a volume of air passing through.

There are two types of friction: supraglottal and glottal friction. Supraglottal friction occurs either

in the oral ([s]), nasal ([m]) or pharyngeal ([ħ]) cavity. Glottal friction is found in the production of whispered sounds, which, according to Ladefoged (1971 pp.6,8)¹¹ are produced with "the vocal cords together or narrowed except between the arytenoids." Breathy-voiced sounds, e.g. [z̤], also have glottal friction.

(iii) Voice. Voice results from the periodic vibration of the vocal cords. The rate at which the vocal cords vibrate in this way determines the pitch of the sound produced; the size of the aperture between the vocal cords for any one cycle of vibration determines the loudness of the sound perceived.

2.6.2. Minor auditory phenomena.

Minor auditory phenomena are those which cannot exist without the major ones.

(i) Nasality. This must be distinguished from "nasalization", an articulatory phenomena. To nasalise a sound is to open the nasal cavity so that air from the lungs may leave the speech tract through the mouth and the nose simultaneously. What we hear while air is passing through the nasal chamber is nasality; the nasality of the bilabial nasal [m] is always accompanied by the major auditory feature voice; while that of the unvoiced [m̥] is also always accompanied by another major auditory feature, friction.

(ii) Percussion. To produce percussion, two organs of speech, at least one of them flexible, come together and

11. Ladefoged, P. (1971), op.cit.

part rapidly. In the case of re-iterative percussion, a loose organ of speech hangs in the way of an air-stream, and then "flaps in the breeze", to use Pike's (1944)¹² words.

2.7.0. Continuant and Non-Continuant.

These are two broad categories which are useful in the classification of speech sounds. In the category "continuant", either the oral or nasal cavity, (or both) is active.

2.7.1. Continuants.

Continuant sounds can be either fricative or frictionless; oral or nasal.

(i) Frictionless Continuants. In the case of frictionless continuants, there must be no audible friction either in the supraglottal cavities or at the glottis; they may be accompanied by plain or creaky voice, since the latter does not entail glottal friction. All this means that there must be no stricture of narrow approximation anywhere in the speech tract, and, finally, the air-stream passing through any stricture must not be as strong as that for the comparable fricative.

Oral frictionless continuants have either a median or lateral mode of egress. The dichotomy slit or grooved is not applicable to this class of sounds.

(ii) Fricative Continuants. Like the frictionless

12. Pike, K.L. (1944), op.cit.

continnants, the fricatives may have a median or a lateral mode of egress, for example, [s, z] and [ʃ, ʒ] respectively, from Karanga. As I have said earlier, the dichotomy slit/grooved applies only to the fricative continnants with a median mode of egress. Thus English [θ, ð], the voiceless and voiced dental fricatives, are slit, whereas [s, z], the voiceless and voiced alveolar fricatives, are grooved. Grooved fricatives produce more, higher frequency random noise on spectrograms than slit ones.

2.7.2. Non-continnants.

These articulations are defined negatively as follows: for an appreciable stretch of time, neither the oral nor the nasal cavity is active. These sounds can be divided into two categories: longer and shorter non-continnants. The auditory correlative of longer non-continnants is plosion. Shorter non-continnants, on the other hand, may consist of either a single percussive tap or what we have called re-iterative percussive. Examples are [ɾ], the flapped r- sound, and [r], the rolled r- sound respectively.

(1) Plosives. These sounds fall within the category of longer non-continnants. There are four different stages to the articulation and perception of a plosive.¹³

The first stage is articulatory, that is, it involves the deliberate adjustment of part of the speech tract. In addition to the velic closure which cuts off

13. Arnold, G.F. (1966) "Concerning the theory of plosives" in Jones and Laver (eds.), op.cit.

the nasal cavity, there is an articulatory closure either at the glottis for the glottal stop, or elsewhere in the oral cavity for other plosives. The auditory correlative of the closure stage is this: if the stop is immediately preceded by a continuant especially in the same syllable, the closure 'chops' off that continuant.

The second stage is the stop period. This involves the passage of time. Pressure builds up behind the articulatory closure until it becomes higher than atmospheric (in the case of implosives, the enclosed body of air is rarefied so that its pressure becomes lower than that of atmospheric air). For 'explosives', this is the compression stage and the vocal folds are either wide apart, if the plosive concerned is unvoiced, or vibrating, if the sound is voiced, for example. The spectra¹⁴ of these sounds show that for unvoiced plosives, there is a period of several centi-seconds when there is a blank on the 'voice-bar' at about 250Hz: for voiced plosives, in the case of Karanga, breathy-voiced plosives, there are some striations at that frequency, most clearly visible on narrow band spectrograms.

The third stage is the release stage. This, like the first stage, is also an articulatory stage: it involves the removal of one or both of the closures made during the first stage. The release can be to a wide or a narrow degree of approximation stricture between the organs which have been in articulatory contact. If

14. See Chapter 5, the section on the plosive phonemes of Karanga.

the release is to a wide degree, the following segment will be characterised by any auditory feature but homorganic friction; and if to a narrow degree, then homorganic friction will be heard. In the latter case, an affricate is produced. In the case of unvoiced stops, there may be a time lag between release of closure and the beginning of voicing for the following vowel. This time lag is called "voice onset time" in the phonetics literature.

The fourth and final stage is plosion. This is an auditory phenomenon, something we hear. Some continuant segment follows the plosion. Depending on the phonetic environment, there may be no plosion heard after release. For example, plosives in utterance final position in some varieties of English are not exploded. But, in spite of the widespread references to "unreleased stops", all stops are released. Otherwise how would it be possible to pronounce the vowel in [bænd] of ['hæt bænd] if the tongue is not released at some stage from its alveolar closure for [t], for example? This seems to be a strong argument against the usual practice of telescoping the third and fourth stages above: if we can have release without plosion, then the two must be separate phonetic phenomena.

2.8.0. A Note on Tone

Shona is a tone language in the sense that the same segmental phonemic sequence can have two different lexical meanings if it has two different tone patterns.

Because tone performs this lexical function, the tones used in the language, /-HIGH/ and /+HIGH/, have been set up as distinct phonemes, along with the segmental phonemes, although in phonemic analysis tones are associated with the syllable.

I shall in this thesis maintain a distinction between pitch and tone, following Carnochan (1962)¹⁵:

"The pitches are many, but the tones are few". 'Pitch' as I have said in the previous chapter, refers to an absolute measurement in Hertz, whereas 'tone' is a classificatory term. Pitch depends on the phonetic environment and other variables mentioned earlier. And a number of different absolute pitches may belong to one tone - /+HIGH/ or /-HIGH/.

Carnochan (1964)¹⁵ relates the regular modifications of the tone pattern of isolate words or phrases to the grammatical functions performed by these structures in particular linguistic contexts. The practice of referring phonological phenomena to varying grammatical functions is illuminating. It represents the total repudiation of the fiction of independent levels of analysis advocated by the structuralists of the earlier part of this century, the post-Bloomfieldians. It was Pike's (1947)¹⁶ view that

"When phonological and grammatical facts are mutually dependent, the treatment of phonology without reference to grammar is a concealment of part of a most important

-
15. Carnochan, J. (1962) "Pitch, Tone and Intonation in Igbo." (1964) "Pitch, Tone and Intonation in Yoruba." In Abercrombie et al (eds) In Honour of Daniel Jones, London: Longmans.
16. Pike, K.L. (1947) "Grammatical Prerequisites to Phonemic Analysis." In E.C. Fudge (ed.) Phonology, Penguin, 1973.

set of structural facts pertinent to phonology... An artificial distinction based upon different approaches to a single set of data may result in a sharp artificial dichotomy which has no reality."

And Palmer (1970, p.XV)¹⁷ says "Phonology is the bridge between phonetics and grammar."¹⁸

I should explain that in describing speech events, we all behave as though the organs of speech assumed discrete postures during speech production. This is a necessary fiction created to facilitate the description of what would otherwise be a perplexing complexity. Speech is a continuum, and involves a multitude of complex movements precisely, though largely unconsciously, synchronised by the brain through the medium of the central nervous system. This role of the nerves in speech production prompted Leonard Bloomfield (1933, p.26)¹⁹ to call speech a means of connecting discontinuous nervous systems: "The gap between the bodies of the speaker and hearer - the discontinuity of the two nervous systems - is bridged by the sound waves."

2.9.0. The Acoustic Phase

Speech travels from the mouth of the speaker to the ear of the listener through the air, in the form

17. Palmer, F.R. (ed)(1970) In his Introduction to Prosodic Analysis, London, Oxford University Press.

18. In Chapter 7, I have tried to explain the modifications of lexical tone in terms of the syntagmatic relationships holding between items in sentences.

19. Bloomfield, L. (1933), op.cit.

of sound waves. The acoustic study of speech events is concerned with specifying the physical characteristics of such sound waves. A number of acoustic concepts are used in such studies. Here I shall give at least a working definition of some of these concepts. In the actual description of the sounds of the Karanga dialect in the following chapters, a relationship will always be sought between the articulatory facts and the acoustic features of any one sound or class of sounds.

2.9.1. Some Acoustic concepts

(a) Frequency. The frequency of a sound wave is usually given as the number of complete repetitions (cycles) of variations in pressure occurring in one second. The unit of frequency measurement used in this thesis is the Hertz. If the vocal folds open and close at the rate of one hundred times in one second during the pronunciation of a sound, the frequency of such a sound is 100Hz. Frequency is a physical phenomenon.

(b) Pitch. The pitch of a sound is the perceptual property of the sound that enables a listener to place the sound on a scale from low to high, without considering the acoustic properties. Generally speaking, when a sound goes up in frequency it also goes up in pitch. But, as Ladefoged (1975, p.163)²⁰ says, there is not a one-to-one relationship between increase in frequency and increase in pitch.

20. Ladefoged, P. (1975) A Course in Phonetics, Harcourt, Brave & Jovanovich Int.

(c) Intensity. Intensity is the amount of energy being transmitted through the air at a particular point, e.g. the ear-drum or the microphone. Intensity is related to the amplitude of vibration. If the amplitude of a sound is doubled, the intensity will increase four times; if trebled, the intensity will be increased nine times. So intensity is proportional to the square of the amplitude. O'Connor (1973, p.82)²¹ says "Intensity is a physical measurement and is not to be directly equated with loudness: the latter is a perceptual matter, like pitch, a matter of hearing and judging what we hear." Intensity is measured in decibels.

(d) Loudness. Loudness depends on the size of variation in air pressure. Just as frequency is the acoustic measurement most directly corresponding to pitch, so acoustic intensity is the appropriate measurement corresponding to loudness.

(e) Formant. The sound spectrograph 'analyses' sounds along three dimensions: the duration of a sound is represented along the horizontal axis, and is here measured in centi-seconds.

The intensity is represented by the darkness of the marks made by the stylus: the darker the marks, the more intense the sound. The frequency is represented along the vertical axis. For voiced sounds, especially the non-constrictive sounds such as the vowels, there are some areas of the spectrograms where there are considerable neatly spaced bands of energy. Such energy

21. O'Connor, J.D. (1973) op.cit.

concentrations are characteristic of the sound in question. They are called formants and are related to the shape and size of the resonating supra-glottal cavities involved in the pronunciation of the sound. Formants are specified in terms of their intensities and centre frequencies.

(f) Transition. Transitions are movements in formant frequency corresponding to the change in the configuration of the vocal tract. But they are not merely the incidental accompaniments of the movements that a speaker makes when moving from one sound to the next. Rather, they are perceptual cues for various consonants. Thus the class of labial consonants will be characterised by formant transitions pointing towards a low frequency area, whereas alveolars will have transitions pointing towards high frequencies.

(g) Locus. The locus theory, to quote Delattre (1968, p.201)²²

"describes the direction which a given formant transition must take in order to contribute to the perception of a given consonantal distinctive feature. In this theory, a locus is the frequency point toward which all the transitions that are related to the perception of the same consonantal distinctive feature converge, regardless of the vowel that precedes or follows."

(h) Resonance. Resonance is the phenomenon whereby one body, which has a natural tendency to vibrate at a certain frequency, will build up vibrations with a comparatively large amplitude when it is set in motion by another body which is vibrating at a similar frequency.

22. Delattre, P. (1968) "From Acoustic Cues to Distinctive Features". Phonetica 18, pp.198-230.

The amplitude of the vibrations will increase as the frequency of the driving system more nearly approaches the natural frequency of the resonator.²³

In the description of vowels, for example, the resonators - the bodies being induced to vibrate - are the supra-glottal cavities. The air in these cavities, having its own resonant frequency, will be set into vibration by the complex periodic waves produced by the vocal cords. These sound waves are the input to the resonators, and the vowel quality we perceive is the output.

2.9.2. Conclusion.

The scope of this outline theoretical framework is essentially very limited. It was not, in any case, intended to be exhaustive of the "articulatory possibilities of man". But it may have raised a number of questions which it is not possible to answer at this stage.

It is a declaration of intent. In it, there may be some finer distinctions than one needs, or even finds possible, to make when confronted with raw linguistic data. But if any concepts are still not clear, their application to such data in the following pages will, I hope, cast more light on them.

23. Ladefoged, P. (1962) Elements of Acoustic Phonetics, pp.55-70. University of Chicago Press, Chicago and London.

CHAPTER 3

3.0 DESCRIPTION AND CLASSIFICATION OF THE VOWELS OF KARANGA

The description and classification of vowels is traditionally done using four parameters:-

- (i) the part of the tongue that is raised: front, central or back;
- (ii) the degree of raising: low, mid or high;
- (iii) the position of the lips: spread, rounded;
- (iv) the position of the soft palate: raised to cut of nasal resonance, or lowered.

On the phonetic level, that is, the acoustic and articulatory level, vowels differ from individual to individual within the same language. Thus Ladefoged (1967, pp.75-76)²⁴,

"I have never yet met any group of potential subjects who normally pronounced a set of words containing a number of phonologically different vowels in such a way that competent phoneticians consider that all the vowels in the corresponding lexical items have the same quality."

In other words, the practice of some writers whereby a vowel is said to be "the same as that occurring in the English word two", for example, is extremely unhelpful.

Consequently, the only feasible way out of the problem is to postulate a set of vowel sounds whose quality does not change from speaker to speaker, and to use such vowels as reference points in the description of the vowels of particular languages. Such a set is the

24. Ladefoged, P. (1967) Three Areas of Experimental Phonetics, O.U.P.

Cardinal Vowel System devised by Daniel Jones (1956, p.34, par.140)²⁵.

I propose to use the term "Cardinal Vowels" in a technical sense, following Ladefoged (1967, p.76)²⁶,

"Whenever the term occurs without further qualification it is intended to designate any sound(s) produced by Daniel Jones and stated by him to be cardinal vowel(s), or any sound(s) produced by any other speaker which are considered by competent observers to be equivalent in phonetic quality to the corresponding cardinal vowel(s) produced by Daniel Jones..."

And Daniel Jones (1956)² maintains that "The cardinal vowels cannot be learnt from written descriptions; they should be learnt by oral instruction from a teacher who knows them." The cardinal vowel system consists of eight primary cardinal vowels and ten secondary cardinal vowels. Cardinal vowel No.1, [i]²⁷ "is the sound in which the raising of the tongue is as far forward as possible and as high as possible, consistently with its being a vowel, the lips being spread." Cardinal vowel No.5 [a], "is a sound in which the back of the tongue is lowered as far as possible and retracted as far as possible consistently with the sound being a vowel, and in which the lips are not rounded."

Beginning from No.1, three auditorily equidistant vowel qualities were established (Daniel Jones says they are acoustically equidistant) by gradually lowering the tongue, with the lips remaining spread and

25. Jones, D. (1956a) An Outline of English Phonetics, Heffer & Sons Ltd. 9th Print, 1969.

26. Ladefoged, P. (1967) op.cit.

27. Jones, D. (1956a) op.cit., pp.31-32.

the soft palate raised to cut off nasal resonance. The symbols [e, ε, a] were assigned to this front series of vowels, numbered 2, 3 and 4 respectively. Then the back of the tongue was raised gradually and another three vowel qualities established in which the lips progressively became more and more closely rounded. The symbols [ɔ, o, u] were assigned to the new vowel qualities, numbered respectively 6, 7 and 8. The soft palate also remained in a raised position throughout.

A secondary series of cardinal vowels was set up by reversing the lip positions of the eight primary cardinal vowels. This second series was designated by the symbols [y, ø, œ, œ̃, ʊ, ʌ, ɤ, ɯ], and numbered 9, 10, 11, 12, 13, 14, 15 and 16 respectively. In addition to this set a pair of cardinal vowels has been added, for which the centre of the tongue has the highest degree of raising. The symbols [ɨ] and [ɯ̥] were used to represent these two vowel qualities, the former of which has lip-spreading, while the latter has lip-rounding. This set of eighteen cardinal vowels thus comprises two sub-categories:

the rounded: [y, ø, œ, œ̃, ʊ, ɔ, o, u, ɯ̥], and

the unrounded: [i, e, ε, a, ɐ, ʌ, ɤ, ɯ, ɨ].

It is also common practice in classifying vowels to refer to those in which the front of the tongue has the highest degree of raising as the FRONT vowels: [i, e, ε, a, y, ø, œ, œ̃], and those in which the back of the tongue has the highest degree of raising as the BACK vowels: [a, ɔ, o, u, ɐ, ʌ, ɤ, ɯ] and,

finally, those in which the centre of the tongue is raised as the CENTRAL vowels: [ɜ , ʌ]. All these parameters are relevant to the classification of the vowels of Karanga. But since there are no back unrounded or front rounded vowels in the language, the secondary cardinal vowels are omitted on Figure 1.

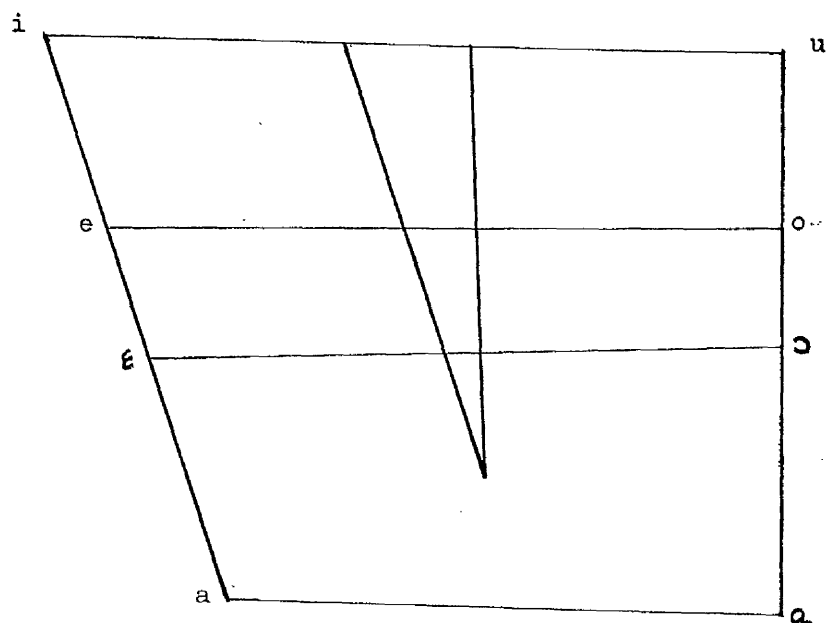


Fig. 1. The Primary Cardinal Vowels.

Karanga (and Central Shona generally) has a total of five vowel phonemes, namely /i, e, a, o, u/. Figure 2 shows the areas in which these vowels are found relative to the eight primary cardinal vowels, indicated thereon by the numbers 1 through 8.

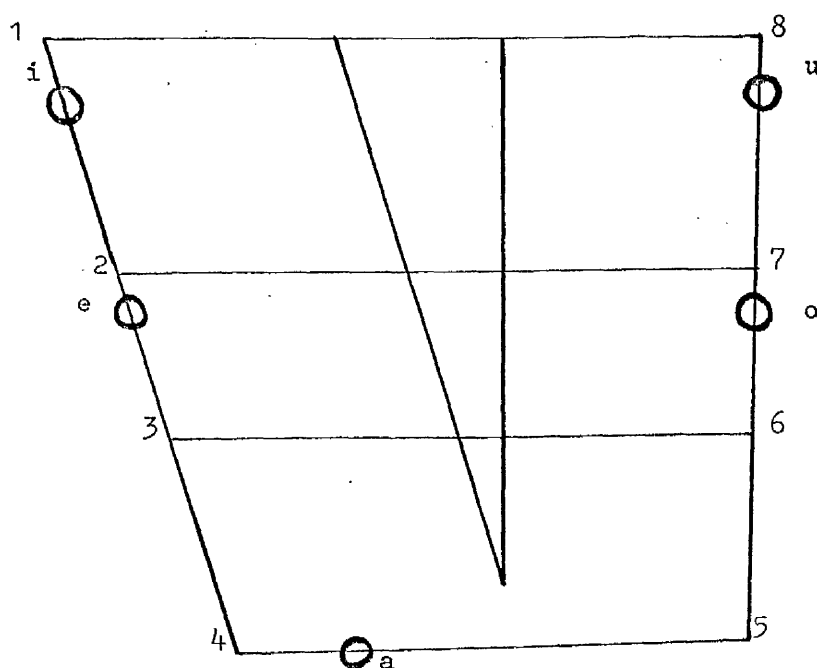


Fig. 2. Vowels of Karanga.

Vowel /i/ is a close front vowel. But it is not as close as Cardinal vowel No.1. It lies in an area between cardinal vowels 1 and 2, being closer to the former than to the latter. The lips are spread and the nasal cavity is inactive throughout, the soft palate being raised.

Vowel /e/ is a front vowel, lying between cardinals 2 and 3. It is nearer to the former in quality than to the latter. This is the so-called MID vowel, as opposed to /i/, which is HIGH, and /a/, which is LOW, for example. The parameters High, Mid, Low refer to the degree of tongue raising.

For /e/, as for all these vowels, nasal resonance is excluded by a velic closure.

Vowel /a/ is neither fully front nor fully back. It lies between cardinal vowels 4 and 5, being

closer to the former than to the latter. The lips are wide open for this vowel. Karanga back vowel /o/ lies between cardinal vowels 6 and 7, but is closer to 7 than to 6. The lips are rounded, but not as closely as they are for the back high vowel /u/, which lies between cardinal vowels 7 and 8. Vowel /u/ is closer to cardinal vowel 8 than to 7 in quality.

As Table II below shows - in column 7 - all these vowels have a long allophone when they occur in a penultimate syllable in an utterance, especially in declarative statements and commands and, more generally, in narrative speech. These vowels can occur after almost all the consonants in the language. In the Karanga dialect I cannot find examples of the naso-oral lateral / nb / followed by the back, close rounded vowel /u/, for example.

3.1. The Vowels of Karanga in Isolation.

According to O'Connor (1973, p.88)²⁸, "in order to characterise vowels in acoustic terms, all we need do is to specify the frequencies and relative intensities of F(ormant) 1, F(ormant) 2, and F(ormant)3." The formant frequencies of the five vowels of Karanga pronounced in isolation by me are shown in Table I(a)²⁹. F1 of the vowel [i] has much greater intensity than F2, while F3 has fairly

28. O'Connor, J.D. (1973) op.cit.

29. These observations are based on the features on spectrograms 1[i] to 5[u].

strong intensity for about three quarters of its duration. The intensity of F3 dies away during the last quarter of the duration of the vowel.

F1 of [e], Spectrogram 2 (Vol.II), has great intensity throughout its duration, while F2 has strong intensity in its middle portion and is weak on either side. F3 is the weakest in intensity.

The F1 of [a] has much more intensity than F2 and F3, which get progressively weaker, that is, F3 is the weakest. F1 and F2 of [o] seem to have equal intensity, while F3 is very weak. Intensity falls progressively from F1 through F3 for [u].

Apart from these observations there are some general patterns among the formant frequencies of these vowels. For example the first formant (F1)'s frequency seems to be closely related to vowel height. Table 1(a) (p.134) shows that F1 increases in frequency as one moves from vowel [i], the front close vowel, to [a], the open low vowel. This formant frequency then starts falling after [a], and is lowest for [u], the back close vowel. The unrounded versus rounded dichotomy between [i] and [u], as well as the front versus back one between the same vowels, seem to have nothing to do with this correlation between vowel height and F1 frequency. The frequency of F1 is therefore inversely proportional to vowel height.

The second point is that the frequency of F2 decreases as one moves from the front close vowel [i] to the back close vowel [u]. Ladefoged (1975, p.173)³⁰ found

30. Ladefoged, P. (1975) op.cit.

that this correlation between the frequency of F2 and the degree of vowel backness was rather unreliable in his analysis of some eight vowels of American English. F2 frequency continued to decrease through the vowels [a, ɔ, ɒ, u] even though some of them are not fully back. This problem does not arise with the vowels which are being analysed here. The frequency of F2 does continue to fall as one moves from the front vowels [i, e] to the non-front [a, o, u]. The fall in F2 frequency between [e], 2.25 KHz, and [a], 1.500 KHz, that is, a difference of 750 Hz, is not as abrupt and steep as that between [a], 1.500 KHz, and [o], 1.00 KHz, that is, a difference of 500 Hz. Obviously [a] is not a front vowel in Karanga, but it is certainly closer to the front series than to the back.

There is a third generalization to be made about these formant frequencies, derived from what has been said above. The distance between the first two formants is greater in the front vowels than it is in back vowels: the more front a vowel is, the greater the distance between F1 and F2 frequency. This evidence corroborates the traditional practice whereby some vowels are shown on the vowel chart as more front than others, even among the so-called front series.

These relationships between formant frequency and vowel height or backness are, of course, not peculiar to Karanga, or for that matter to Central Shona. They are the acoustic correlates of physiological phenomena: the degree of tongue raising, whether back or front, has a direct bearing on the shape and size of the supra-glottal cavities contributing to the quality of the vowel

sound produced through resonance. In an article published in Word, Delattre et al. (1952, pp.195-210)³¹ reported the results of an experiment in which they synthesised sixteen vowels which they called "cardinal vowels", though not in the technical sense adopted here. They seem to think that the phonetic quality of the cardinal vowels will differ depending on the native language of the speaker. A vowel that varies in that way is, for us, not a cardinal vowel.

However, the point of interest to emerge from their investigation is that the correlations between formant frequency and degree of vowel backness and vowel height (their Fig. 1, p.199) are confirmed by their synthetic sounds. They also conclude that only the first two formants are, apparently, essential for the identification of individual vowel qualities, while the higher formants seem to serve as cues for identifying particular speakers. O'Connor (1973, p.87)³² concurs.

As for the accuracy of the observations based on the spectrograms in this chapter and others following, one need only repeat what other people have said about the interpretation of spectrograms. Within reasonable limits, there are differences of emphasis from one researcher to another, a fact which, to quote Ladefoged (1975)³³ makes the reading of spectrograms "more an art

31. Delattre, Liberman, Cooper and Gerstman (1952), "Experimental Study of the Acoustic Determinants of Vowel Color." Word, 8.

32. O'Connor, J.D. (1973) op.cit.

33. Ladefoged, P. (1975) op.cit.

than a pure science." What various acoustic phoneticians will, it is claimed, confirm are the correlations which I have observed in the vowels above. The exact values of the formant frequencies in Hertz, if slightly different, need not detract from the importance of these general features. Such differences are minimised by pronouncing the examples used several times and at different pitches in order to see the average general tendency.

Fant (1970)³⁴ remarks, "The number of segments found in a spectrogram of an utterance will vary according to the investigator's ambitions."

3.2. The hy-complexes

As we have seen in the first chapter, "depression" or "depressor" is a term derived from the lowering effect which breathy-voiced or murmured consonants have on the pitch of the syllable in which they occur as the initial segment. The exact nature of the vocal fold configuration during the production of these consonants has eluded many researchers. But in this thesis I shall assume that Ladefoged (1971, pp.6,8)³⁵ is a fair analysis of the phenomenon. Ladefoged has done some work using Shona informants and others from the Bantu area, as reported in that book.

34. Fant, G (1970) "Analysis and Synthesis of Speech Processes" In Malmberg (ed.) op.cit.

35. Ladefoged, P. (1971) op.cit.

The glottal fricative / $\underset{\cdot}{h}$ / is the "depressor" (here the "breathy-voiced") consonant par excellence. It is customary for some writers to say that breathy-voiced consonants have an $\underset{\cdot}{h}$ -quality, meaning that there is a guttural quality about such consonants. This glottal fricative has no plain-voiced or voiceless counterpart in Karanga. But it combines only with the vowels of the dialect - also of Central Shona - in what I have here called the $\underset{\cdot}{h}V$ -complexes, where [$\underset{\cdot}{V}$] stands for any of the five vowel sounds discussed above.

The ensuing analysis of the $\underset{\cdot}{h}V$ -complexes is expected to throw some light on the nature of the relationship between the vowel and its breathy-voiced glottal onset. All the examples were recorded pronounced by me at the same pitch and then analysed using the sound spectrograph - Kay Elemetrics model - on a broad band filter (Spectrograms 6 through 10). And amplitude sections were made at points at which it seemed that the articulating and phonating organs had attained their steady state for $\underset{\cdot}{h}$ - and for $\underset{\cdot}{V}$.

The notation $\underset{\cdot}{h}V$ is intended to reflect the fact that the fricative and the vowel are being regarded as a sequence. The sections show two different profiles for each of these segments. However, when we come to the so-called breathy-voiced consonants, a different relationship will be postulated between the consonant and this breathy quality. From an auditory point of view, the breathiness and the consonant are simultaneous. This is a phonetic judgment. There might be grounds for a sequential analysis of [$\underset{\cdot}{b}$] as [$\underset{\cdot}{b}\underset{\cdot}{h}$], [$\underset{\cdot}{m}$] as [$\underset{\cdot}{m}\underset{\cdot}{h}$] for

example, from a phonological point of view, even though phonetically the two are co-extensive.

However, a number of interesting similarities and differences stand out on the $\underset{\cdot}{h}\underset{\cdot}{V}$ spectrograms, summarised on Table 1(b):

- (i) the formant structure of vowels has, on the whole, bolder outlines here than when the vowels are in isolation;
- (ii) their formant frequencies largely conform to those of the vowels in isolation.³⁶
- (iii) $[\underset{\cdot}{h}]$ has a clear formant structure, not radically different from that of the associated vowel, especially for the first three formants;
- (iv) $[\underset{\cdot}{h}]$ has very little energy above 3 KHz;

The following differences between $[\underset{\cdot}{h}]$ and $[\underset{\cdot}{V}]$ are also apparent:

- (i) the striations of $[\underset{\cdot}{h}]$ are concentrated in the very low frequencies and to about three quarters of the latter part of the segment. During the first part of the segment the vocal cords presumably have not attained their steady state for the fricative.
- (ii) The striations are farther apart for $[\underset{\cdot}{h}]$ than for the adjacent vowels, which means
- (iii) $[\underset{\cdot}{h}]$ will have lower pitch than the vowel.
- (iv) The centre frequencies of the formants of $[\underset{\cdot}{h}]$ are lower than those of the ensuing vowel. Consequently,
- (v) There are some very short formant transitions separating $[\underset{\cdot}{h}]$ from $[\underset{\cdot}{V}]$. These transitions are

36. Fig. 3, p.135. is a graph on which F1 and F2 of the vowels have been plotted against F1 and F2 of $[\underset{\cdot}{V}]$.

negligible compared to those between the two elements of a vowel sequence (see Spectrograms 11 through 29 and the next section). The close resemblance between the amplitude sections of [h̥] and that of [V̥] means that the vocal tract configuration for the one is not too different from that for the other.

- (vi) In all the five examples, the vowel has greater duration than [h̥]: the boundary between the two being taken to be the point at which striations, wide apart for h, begin to be close together. The formant transitions, infinitesimal though they are, also help to show the location of this divide.
- (vii) The vowel has higher amplitude peaks than the fricative, although
- (viii) These amplitude peaks occur at corresponding points along the horizontal frequency axis for [h̥] and for [V̥].

Now, the extent to which the sound waves from the periodically vibrating vocal cords are amplified by the supra-glottal resonance chambers depends on how nearly the harmonics of these sound waves - including the fundamental - approximate to the resonant frequency of the chambers. The shape and size of these chambers are different for different vowels, hence the differently shaped vowel spectra and the diverse qualities of the vowels perceived. So the spectral similarity between [h̥] and [V̥] suggests, as we have said, that the two have close to identical vocal tract configurations. Since they differ mainly in the size of their amplitude

peaks, I conclude that they must have different fundamental frequencies and different harmonic compositions.

The fundamental frequency of [h̥] must be lower than that of the vowel, a conclusion corroborated by the greater distance between the striations of [h̥]. This difference between the fundamental frequency of the two segments also, logically, corresponds to a pitch difference between them.

That [h̥], the Voiceless glottal fricative, or [h̥], its breathy-voiced counterpart, is merely the onset of the ensuing vowel, its voiceless or breathy-voiced version, may be what these spectrograms and amplitude sections are confirming. C.M. Doke (1931, p.92) ³⁷ says the Shona [h̥],

"might be described as a strong throat-roughening of the vowel, 'ha' differing from 'a' in that roughening, for the effect of the voiced throat friction is observable throughout the duration of the vowel."

I agree with Doke to the extent that the first part of this quotation suggests that [h̥] is a modification or version, ("a throat-roughening") of the ensuing vowel. There is evidence on the spectrograms 6 through 10 to suggest that [h̥] affects [v] qualitatively, and pitch-wise: [h̥] being on a lower pitch than [v], there is an upward pitch movement from the end of [h̥] to the steady-state of [v]. That the first part of [v] has lower pitch than its centre is an assimilatory matter, after the lower pitched glottal fricative.

37. Doke, C.M. (1931) op.cit.

What seems to be beyond doubt is that the complex [h̥-V] is, auditorily and acoustically, like two juxtaposed allophones of the same vowel, recognised as the "same" by their spectral similarity. J.D. O'Connor (1973, pp.78-79)³⁸ says that "The same vowel may be said at different pitches, that is, with different fundamentals. Consequently, the harmonic structure will be different."

[On Table 1(a) formants 4 and 5 were given in the hope that they might throw some light on the acoustic character of the vowels. In the event, the first 3 formants were found to have all the information required in this exercise.

The graph below (Fig. 3) is interesting for a number of reasons: (a) it shows the close correspondence between instrumentally based descriptions and traditional plottings on the vowel quadrilateral which are based on articulatory perceptual criteria, here for both plain and breathy-voiced vowels; (b) the centre frequency of the F2 and F3 of [V̥] in the context of the glottal fricative is consistently lower than that of the formants of the plain vowels.]

38. O'Connor, J.D. (1973) op.cit.

Interpretation of Vowel Spectrograms

The vowels in isolation

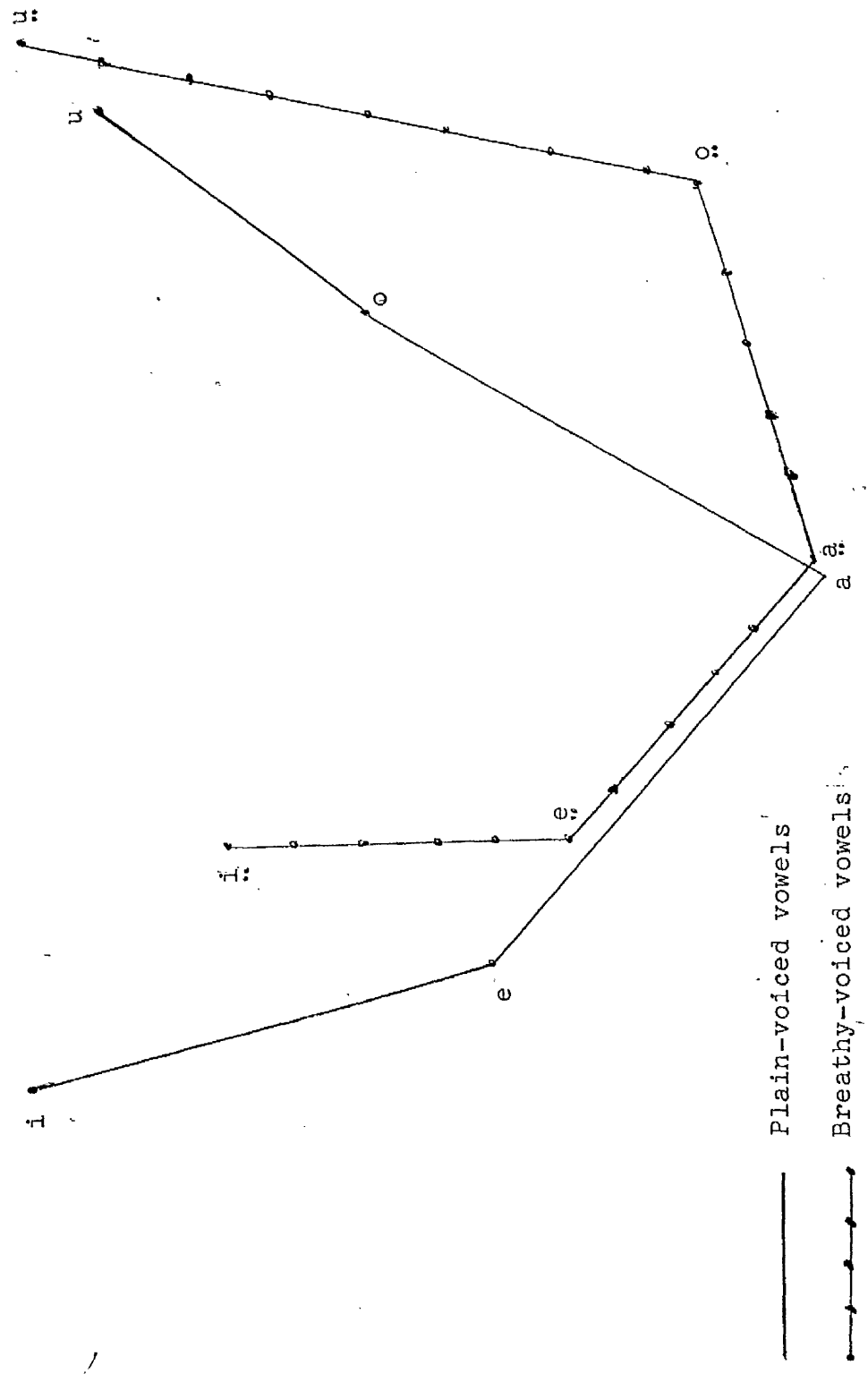
TABLE 1(a)

	Vowel	Duration in centi- seconds	F1	F2	F3	F4	F5
1.	i	56	225Hz	2.5K	3.5K	4.7K	5.50K
2.	e	56	400Hz	2.25K	3K	3.5K	5.50K
3.	a	56	725Hz	1.5K	2.5K	3.50K	5.00K
4.	o	60	350Hz	1K	2K	2.5K	6.00K
5.	u	60	250Hz	725Hz	2.25K	3.5K	-

TABLE 1(b) Spectrograms of hV-complexes.

		Duration	F1(h)	F2(h)	F3(h)
6.	hi	h.... 19c-s i.... 30c-s	300	2.20K	3.00K
7.	he	h... 7c-s e... 45c-s	430	2.00K	3.50K
8.	ha	h... 13c-s a... 45c-s	520	1.45K	2.30K
9.	ho	h... 8c-s o... 50c-s	470	530	2.45K
10.	hu	h... 6c-s u... 50c-s	220	525	-

200Hz
250Hz
300Hz
400Hz
500Hz



Plain-voiced vowels

Breathy-voiced vowels

Fig. 3

For example:

<u>Vowel</u>	<u>Fundamental</u>	<u>2nd Harmonic</u>	<u>3rd Harmonic</u>
[a]	{ 100 Hz	200 Hz	300 Hz
	{ 175 Hz	350 Hz	525 Hz

O'Connor continues,

"The vowel will be recognised as the same in both cases because the general shapes of the two spectra will be the same, i.e., the harmonics with the greatest amplitude will be at the same frequency in both, regardless of what number the harmonics have in the structure. The general shape of the spectrum characterises a particular vowel, rather than the actual number and frequencies of individual harmonics."

The case described so clearly by O'Connor concerns the recognition of a vowel as vowel [x] when it is pronounced at different pitches by different people or by the same person on different occasions. Our [h̥V] complexes are like such a vowel, the only difference being that here the two realizations are juxtaposed. As in the case described by O'Connor, there is no difference in the vocal tract configuration for [h̥] and for [V]: that seems to be identical. The difference between the two seems to arise from the state of the glottis for each: there is a change from its "posture" for [h̥] to another for the vowel, with the latter more ideal for voice production than the former. The short formant transition from [h̥] to the vowel, in other words, reflects more the increase in fundamental frequency

from 100 Hz for [h̥] to 175 Hz for [ɣ], for example, than a change in the shape and size of the supra-glottal cavity.

The relationship between the amplitude sections of [h̥] and [ɣ], that is, having peaks of energy at the same points along the horizontal, frequency axis, is paralleled by that which exists between the sections of homorganic, voiceless and breathy-voiced fricatives (see Spectrograms 63-75). In these examples, as we know, the supra-glottal cavity shape and size are identical for each pair; only the configuration of the vocal cords is different. The voiceless fricative has greater noise amplitude - since the wide-apart vocal cords allow a big volume of aperiodically vibrating air through - than its breathy-voiced counterpart, when the vocal cords are in the position described by Ladefoged (1971, pp.6,8)³⁹ for breathy-voice or murmur. In the same way the vowel has greater voice amplitude, because the vocal cords are ideally positioned for voice production, than its counterpart, breathy-voiced [h̥], or its 'reduced' pre-posed self.

3.2.1. The Vowel 'Sequences'

The five vowels of Karanga described above occur singly as syllable nuclei; or they can be found in sequences of two, sometimes three. The point of interest lies in the latter type of occurrence: how are such vowel sequences to be interpreted? Do they constitute a single

39. Ladefoged, P. (1971) op.cit.

complex syllable nucleus or are they single nuclei of two (or three) abutting syllables? If the former is the case then we would have to recognise a series of diphthongal phonemes such as English /ai/ cf /¹tai¹m/, "time", for example.

What criteria are we going to use in reaching a decision one way or the other? Are there any phonetic features that can help us or are we going to use grammatical criteria. I propose to deal with the problem as follows:

(a) to describe the sequences in phonetic, mainly acoustic, terms and see if any answer will emerge on that level.

The point is, we already know⁴⁰ the formant frequencies of the vowels involved in the sequences, so we have an idea of the direction and extent of the corresponding formant 'transitions' between juxtaposed vowels;

(b) to consider the grammatical function of the vowels and see if the treatment suggested by such grammatical function coincides with our phonetic answer, if any.

I shall divide the vowel sequences into sub-systems as follows:

- (i) those that can be divided into two syllables corresponding to two grammatical formatives;
- (ii) those that can be divided into two syllables but are assignable to one grammatical/ lexical item;

40. That the formant frequencies of the vowels in Spectrograms 11-29 (the sequences) turn out to be the same as those in which the vowels were pronounced in isolation and at a different pitch is helpful. The formant frequencies recorded on Tables I(a) and I(b), and Tables II(a)-(c) are therefore the averages of several instances or pronunciations of the vowels.

- (iii) those between which an optional j-glide can be inserted;
- (iv) those between which an optional w-glide can be inserted;
- (v) those between which no glide can be inserted at all.

3.2.2. Description of the Vowel Sequences

This description of the vowel sequences is supplemented by a summary of some of the outstanding acoustic features observable on the spectrograms of the vowel sequences - Table II, pp.140-142

Utterances containing the vowel sequences were recorded in a sound proofed studio, pronounced by me. Each utterance was then analysed using the sound spectrograph. The accompanying Table II (a)-(c) gives:

- (i) the vowel sequence in question in column 1;
- (ii) the formant relationships holding between the two vowels, in column 2;
- (iii) an utterance in which the vowels occur, in column 3;
- (iv) the duration, in centi-seconds, of the first vowel, V_1 , in column 4; the duration, that is, of both the initial transition and the steady state of V_1 ;
- (v) the duration of the terminal transition (F2) of V_1 , i.e. the formant frequency movement before reaching the steady-state of the second vowel, V_2 , in column 5;
- (vi) the duration of the transition in (v) above and the duration of the steady-state of V_2 , in column 6; (see pp.168-169 for explanation);

TABLE II (a)

1 Vowel sequence	2 Formant Relationship			3 Utterance	4 Duration of V ₁	5 Duration of Formant Transition	6 Duration of V ₂	7 Duration of V in - CV#
	F ₁	F ₂	F ₃					
11. ae	a=725Hz e=400	1.5K 2.25K	2.5K 3K	taéndà	15	10	30	40
12. ai	a=725Hz i=225Hz	1.5K 2.5K	2.5K 3.5K	taímbá	14	12	25	37
13. ao	a=725Hz o=350Hz	1.5K 1K	2.5K 2K	taóná	13	9	25	34
14. au	a=725Hz u=250Hz	1.5K 725Hz	2.5K 2.25K	taúra	15	8	29	37
15. ei	e=400Hz i=225Hz	2.25K 2.5K	3K 3.5K	sei	23	5	18	28
16. eu	e=400Hz u=250Hz	2.25K 725Hz	3K 2.25K	seúra	16	11	26	37

1 Vowel sequence	2 Formant Relationship			3 Utterance	TABLE II(b) 4 Duration of V ₁		5 Duration of Formant Transition	6 Duration of V ₂	7 Duration of V in - CV#
	F1	F2	F3						
17. ea	e=400Hz	2.25K	3K	se`angú	13		10	23	33
	a=725Hz	1.5K	2.5K						
18. ia	i=225Hz	2.5K	3.5K	ndi`ani	16		8	28	36
	a=725Hz	1.5K	2.5K						
19. ie	i=225Hz	2.5K	3.5K	ndi`e	26		5	15	31
	e=400Hz	2.25K	3K						
20. io	i=225Hz	2.5K	3.5K	ndi`one	11		5	29	31
	o=350Hz	1K	2K						
21. iu	i=225Hz	2.5K	3.5K	ndi`umbe	13		not clear	28	28
	u=250Hz	725Hz	2.25K						
22. oi	o=350Hz	1K	2K	toi`nda	11		7	23	30
	i=225Hz	2.5K	3.5K						

1 Vowel sequence	2 Formant Relationship F1 F2 F3	3 Utterance	4 Duration of V ₁	5 Duration of Formant Transition	6 Duration of V ₂	7 Duration of V in-CV#
23. oe	o=350Hz 1K 2K e=400Hz 2.25K 3K	toerera	10	5	27	30
24. oa	o=350Hz 1K 2K a=725Hz 1.5K 2.5K	toàtorà	13	5	15	27
25. ou	o=350Hz 1K 2K u=250Hz 725Hz 2.25K	toùtorà	8	not clear	18	30
26. ui	u=250Hz 725Hz 2.25K i=225Hz 2.5K 3.5K	muímbe	14	5	23	28
27. ua	u=250Hz 725Hz 2.25K a=725Hz 1.5K 2.5K	muàtorè	10	5	16	29
28. ue	u=250Hz 725Hz 2.25K e=400Hz 2.25K 3K	muéndè	6	5	30	35
29. uo	u=250Hz 725Hz 2.25K o=350 1K 2K	muòndé	15	not clear	27	27

(vii) the duration of the vowel preceding the last vowel in the utterance is given in column 7. In a number of these examples the vowel preceding the last vowel of the utterance is also the second vowel, V_2 , of the sequence. In such examples the duration of V_2 is identical to that of the vowel in utterance penultimate position, i.e. - CV#.

This last column gives information relevant to the phenomenon in Shona phonology, whereby the vowel in the penultimate syllable is lengthened. The utterances used in this analysis comprise declarative statements, interrogatives, noun phrases as well as commands. If V_2 in our sequences does belong to a separate syllable, then column 7 will have provided useful phonological information about penultimate syllable lengthening, information that will be used as evidence for the argument that the two members of a vowel sequence belong to different vowels and are not a diphthong.

3.2.3. Interpretation of the Spectrograms of the Vowel Sequences

As I have explained above, formant transitions are the acoustic correlates of the changes in the vocal tract configuration (changes) made by the organs of speech moving from one articulatory "posture" to the next. The vowel sequences under consideration here comprise vowels of different qualities, and consequently the transitions between the vowels are generally very pronounced. Here I shall concentrate on the vowel

formants in so far as they give information about the sequences. Information given by initial and terminal transitions about preceding or following consonants, although valuable in another context, will be noted only in passing.

Spectrogram 11 is of the utterance [tãénda] "we have gone". There is an initial F2 transition of the vowel [a] after the consonant [t]. This transition has a locus at about 1.8 KHz. It reaches the steady-state of [a] at about 1.25 KHz. Both the transition and the steady-state of F2 have a duration of about 15 centi-seconds. After this, there is a fairly steep transition, without a break intervening, involving both F1 and F2 of [a] and lasting for 10 centi-seconds. F1 falls in frequency to reach the steady-state of the corresponding formant of V₂, [e]; while F2 of [a] rises to reach the steady-state of the corresponding formant of V₂, [e]. This second vowel has a steady-state duration of 30 centi-seconds which, together with the terminal transition of [a], makes a total duration of 40 centi-seconds.

This second vowel [e] has the greatest duration of all the segments in the utterance: it is also the penultimate vowel.

Spectrogram 12 [tãímbá], "we have sung", also has some interesting features. The initial F2 transition of the vowel [a] has much the same locus as in the previous example, since the vowel is here preceded by the same consonant [t]. The duration of [a] is also the same as before. But the terminal transitions from the

142

steady-state of both F1 and F2 of [a] to the steady-state of the corresponding formants of V_2 , namely [i], are steeper and go farther, compared to example 11. F2 rises while F1 falls in frequency to reach the steady-state of the appropriate formants of V_2 . The transition lasts for 12 centi-seconds; the steady-state of [i], which is also the penultimate vowel of the utterance, lasts for 25 centi-seconds which, together with the terminal transition of [a], makes a total duration of 37 centi-seconds. There is again no break on the spectrogram corresponding to the boundary between the two elements of this sequence.

Spectrogram 13 is of the utterance [təóná], "we have seen". The initial F2 transition is also similar to that of the two examples above. The duration of both this transition and the steady-state of the vowel [a] is 13 centi-seconds. In this example, as the table showing formant relationships suggests (p.140), both F1 and F2 of [a] fall in frequency to reach the steady-state of the corresponding formants of V_2 , here [o]. The transition has a duration of 9 centi-seconds, and, added to the duration of the steady-state of V_2 , [o], (25 centi-seconds), gives a total duration of 34 centi-seconds. The V_2 here is also the penultimate vowel of the utterance.

Spectrogram 14 is of the utterance [táúrá], "speak!" The initial F2 transition of [a] after [t] is again similar to that in previous examples. The terminal transitions of F1 and F2 of this vowel fall in frequency to an even lower level than in the example 13 above. This is to be expected since [u], our V_2 here, has lower centre frequencies for F1 and F2 (250 Hz and 725 Hz

respectively) than V_2 above, [o], (350 Hz and 1.00 K respectively).

The vowel [a] has a duration of 15 centi-seconds, including its initial F2 transition; while [u] including the terminal transition of [a], has a duration of 37 centi-seconds. The latter vowel is also the penultimate vowel of the utterance, and has the greatest duration of all the segments in the utterance. On this spectrogram its higher formants, F3 and F4, have very little energy.

Spectrogram 15 is of the utterance [séi] "how" or "why"? The fricative [s] has its noise component in the high frequencies above 4.5 KHz. To that extent it is similar to [t], whose noise burst is also concentrated in these high frequencies. But the F2 transition of the vowel [e] after [s] points towards a lower locus than did the transition of the corresponding formant of [a] after [t]. Here the locus is as low as 1.00 KHz. This suggests a difference in place of articulation between [t] and [s].⁴¹

However, the vowel [e] has the greatest duration of all the segments in the utterance, 28 centi-seconds. Vowel [i], the V_2 , has a duration of 18 centi-seconds. The former vowel has its striations closer together than those of [i]; which is to be expected: [e] is on a high tone (and therefore has higher pitch) while [i] bears low tone and is in utterance final position.

41. This is in fact the case, as a comparison of Palatograms (i) a and b and Palatograms (ix) a and b shows: [t] is an alveolar sound, whereas [s] is dental.

Generally speaking, this gives us perceptually different allophones of the same vowel depending on whether they are the nuclei of a high tone or a low tone syllable, and also on whether the syllable occurs in the middle or at the end of an utterance.

Spectrogram 16 is of the utterance [sèúrá], "weed", clear an area of grass". The noise component of [s] is in the same high frequencies as in 15 above. The initial F2 and F3 frequency transitions of the vowel [e] also point towards low loci (compared to their loci after [t]) of between 1.00 KHz and 1.50 KHz. Both F2 and F3 of this vowel seem to have a shorter duration, than F1, about 17 centi-seconds and 23 centi-seconds respectively.

Only F2 falls in frequency to reach the steady-state of the corresponding formant of V₂, [u]. This transition is rather steep and lasts about 11 centi-seconds. The third formant of [e] rises in frequency to reach F3 of [u] at about 3.30 KHz. Apart from F3, which is only barely visible, [u] does not have higher frequency formants than F2 on this spectrogram. This vowel has a duration of 37 centi-seconds. The terminal transition of F2 of [u] before [r] points towards a locus of about 2 KHz.

On spectrogram 17, [séàngú], "like mine", class 6, the initial F2 transition of [e] after [s] points towards a low locus, as in the previous two examples. Both the second and third formants of [e] fall in frequency to reach the steady-state of the corresponding formants of [a]. But F1 of [e] rises in frequency to

reach the steady-state of the first formant of [a].

The transitions last about 10 centi-seconds. And the steady-state of the penultimate vowel has a duration of 23 centi-seconds which, together with the transition from [e] comes to 33 centi-seconds. The terminal transitions of F1 and F2 of the vowel [a] before [ng] seem to point towards a common locus: a defining characteristic of all velar-consonant articulations.

Spectrogram 18 [ndiáni], "who is it?" is outstanding for the clearness of the initial frequency transitions of F2 and F3 of the vowel [i] preceded by the naso-oral alveolar complex [nd]. F2 of this vowel points towards a locus of about 1.80 KHz, suggesting that this naso-oral sound has the same place of articulation as the voiceless plosive [t] discussed briefly above. This holds for all the naso-oral complexes in examples 18 through 22.

The transitions of these two formants are abrupt, and both F2 and F3 have a rather short steady-state. The vowel as a whole has a duration of 16 centi-seconds. The transition from [i] to [a] lasts 8 centi-seconds, and both F2 and F3 fall in frequency to reach the corresponding formants of [a]. The first formant of [i], like that of [e] in example 17 above, rises in frequency to reach the steady state of [a]. The latter vowel has a duration of 28 centi-seconds in its steady-state. Together with the on-glide from [i], the vowel has a duration of 36 centi-seconds.

Spectrogram 19 [ndi^é], "it's him" or "he's the one", is remarkably clear. The second formant of [i] points towards a locus of about 1.8 KHz after [nd]. The transition to the steady-state of [i] is accomplished very quickly, and this vowel, being in a penultimate position, has the greatest duration of all the segments in the utterance, 31 centi-seconds.

As in example 15 above, the final vowel in this example is on a low tone. Consequently, its striations are farther apart, showing allophony between [é] and [è], and between [i'] and [i].

Spectrogram 20 is of the utterance [ndi^{ón}è], "let me see". The initial F2 and F3 frequency transitions of [i] preceded by [nd] are identical to those in examples 18 and 19. The vowel has a duration of 11 centi-seconds. The second formant transition of [i] before [o], the second vowel, is abrupt, falling steeply in frequency in order to reach the corresponding formant in the second vowel. The transition lasts about 5 centi-seconds. There is no apparent change in frequency between the first formants of both vowels. The latter vowel [o] has clear F4 and F5, and a duration of 34 centi-seconds, comprising the duration of its steady-state (26 centi-seconds) and that of the on- and off-glide (5 centi-seconds and 3 centi-seconds respectively).

Just as the initial second formant frequency transition of [i] preceded by [nd] points towards a locus at about 1.8 KHz, so does the second formant transition of [o] before the nasal consonant [n].

Spectrogram 21 is of the utterance [ndi^{úm}bè],

150

"let me mould". There are no clear formant transitions between V_1 and V_2 . V_1 [i], has an initial F2 transition with a locus of about 1.80 KHz. Initial F2 and F3 transitions of [e] preceded by [mb] point towards low loci. The second formant transition points towards a locus of 1.00 KHz, that is, the same locus as for the fricative [s] in examples 15, 16 and 17 above.

The first vowel [i], has a duration of 13 centi-seconds; whereas the second vowel [u], which is also the penultimate vowel, has a duration of 28 centi-seconds.

On Spectrogram 22, [t'óíndà], "we are going, about to go", the initial transition of F2 after [t] is short and abrupt, only about 3 centi-seconds. The vowel [o] as a whole has a duration of 11 centi-seconds. The terminal F2 transition of [o] before the second vowel [i] is also very short, about 7 centi-seconds. The second vowel has a duration of 30 centi-seconds.⁴²

On Spectrogram 23 [t'óérèrà], "we are floating away", both the second and third formants of [o] have terminal transitions to the corresponding formants of [e], the second vowel. [o] has a duration of 10 centi-seconds, and the transition of F2 lasts about 5 centi-seconds. The second vowel in this sequence is not the penultimate vowel of the utterance and has a duration of 27 centi-seconds; whereas the vowel [e] which is in

42. On this spectrogram, because the naso-oral complex is in an intervocalic position, the nasal element has some very clear formants, with transitions before the [-d] which point towards the same loci as do those of the vowel [a] after [-d].

penultimate position has a duration of 30 centi-seconds. The terminal F2 transition of the vowel [e] before [r], the rolled consonant, points towards a 1.25 KHz locus.

Spectrogram 24 is of the utterance [tóàtórá], "we are going to take them", class 6. The most interesting feature of this spectrogram is the way in which the formants of the first vowel have paired up with the corresponding formants of the second vowel, forming a symmetrical pattern:

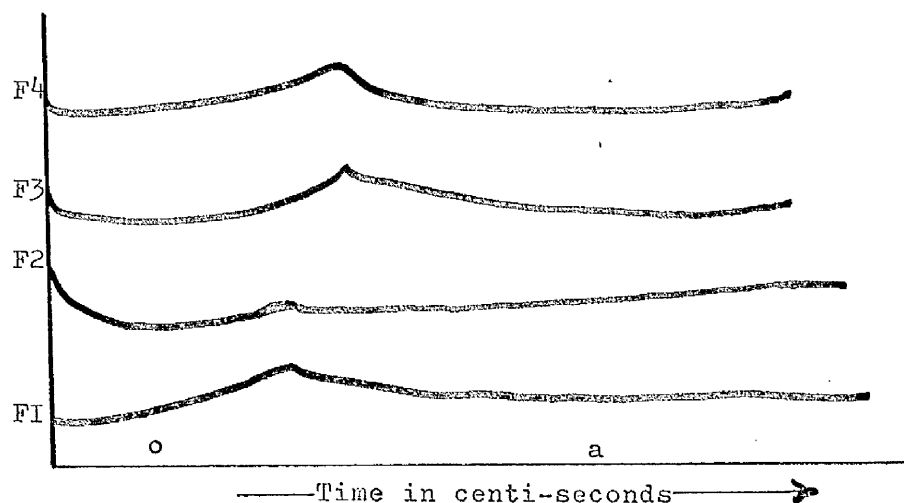


Fig. 4. The formant pattern of [-oâ-]

As Figure 4 shows, this patterning arises from the continuity of the first into the second vowel, a fact which holds true for all this sequence. Both vowels in this particular sequence seem to have equal intensity for corresponding formants. The first vowel has a duration of 13 centi-seconds, while the second, which, as in example 23 above, is not the penultimate vowel, has a duration of 20 centi-seconds. The vowel in the penultimate syllable in this utterance has a duration of 27 centi-seconds.

On spectrogram 25 [tòùtórá], "we are about to/going to take it", class 3, the formants of the two juxtaposed vowels form a pattern similar to that of the pair in example 24.

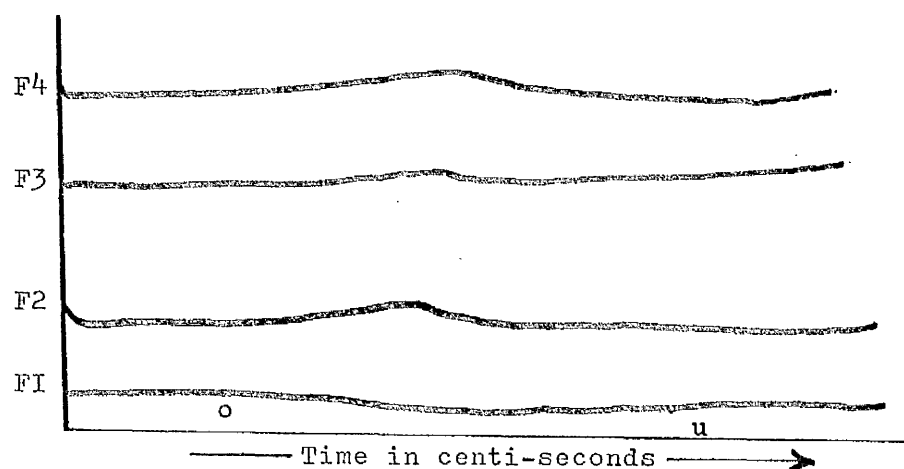


Fig. 5. Formant pattern of [-òù-]

The only difference is that on this example only F3 and F4 of each vowel display this symmetry clearly. The vowel [o], V_1 of the sequence, has a duration of 8 centi-seconds. The formant transitions between [o] and [u] are infinitesimal. The second vowel, [u] has a duration of 18 centi-seconds, compared to 30 centi-seconds for the vowel in the penultimate syllable of the utterance.

On Spectrogram 26, [múímbe], "you (pl.) must sing", the second formant of [u] has a steep frequency transition to the corresponding formant of the following vowel, [i]. Since [i] is the penultimate vowel in the utterance, it has the greatest duration of all the segments in this utterance, 28 centi-seconds, including 5 centi-seconds for the transition from the preceding

vowel.

Note that the last vowel in this utterance, [e], has low tone and, consequently, the striations are farther apart than those of other vowels with high tone and, therefore, higher pitch. This particular vowel also dies away quite suddenly. The second nasal on this spectrogram has F2 and F3 frequency transitions pointing towards very low loci before [-b]: they have the same low loci as the initial transitions of the corresponding formants of the last vowel, [e], preceded by the same sound [-b-].

On Spectrogram 27, [muàtóre], "you (pl.) must take them" class 6, the vowel [u] seems to have very short duration, about 9 centi-seconds. The transition of its F2 to the corresponding formant of the second vowel lasts about 5 centi-seconds. Although F1 of [a] and F2 of [u] seem to have the same centre frequency, it is remarkable how the latter (F2 of [u]) by-passes F1 of [a] to connect with the corresponding formant of [a]. The latter vowel has a duration of 21 centi-seconds, compared to 29 centi-seconds for the vowel in the penultimate syllable of the utterance.

On Spectrogram 28, [muéndè], "you (pl.) must go", the vowel [u] again seems to have a very short duration, about 6 centi-seconds. This is 2 centi-seconds shorter than the duration of the transition of its F2 to the corresponding formant of the second vowel, [e]. Again, although F2 of [u] is closer in frequency to F1 of [e], it nevertheless by-passes the latter and rises in frequency to connect with the corresponding formant of [e]. This

also happens in examples 11, [ae]; 22, [oi]; and 27,[ua] above.

The second vowel in this sequence [e], has very clear formants and has a long duration, 38 centi-seconds, since it is in the penultimate position in the utterance.

Although the vowel [u] has very little energy on Spectrogram 29, [mùòndé], "fig-tree", it has fairly long duration, 15 centi-seconds. There are no clear formant transitions between the vowels [u] and [o], perhaps because the vocal tract configurations for these vowels are not too radically different.

This spectrogram displays the major acoustic features observed on all the examples above: there is no break between the two vowels in the sequence; the vowel in the penultimate position in the utterance has the longest duration, 24 centi-seconds. The final vowel of the utterance, [e], is on a high tone and so its striations are closer together (for higher pitch) than those of the 'same' vowel in example 28 above.

In all the examples described above, the second vowel of the sequence has longer duration than the first. This is an important observation since it can be used as evidence in support of the argument that the vowels in the sequence in question belong to different syllables. Since in the language, the vowel in the penultimate syllable in an utterance has longer duration than other vowels, we can conclude that the second vowel in these sequences, where it is in a penultimate position, belongs to a separate vowel. In examples 11 through 22,

and 26 through 29, the second vowel is, according to this reasoning, the nucleus of a separate, in this case, penultimate syllable.

However, in examples 23, 24 and 25, where the second vowel of the sequence is in an ante-penultimate position, it (the second vowel) is nevertheless longer in duration than the first vowel. The conclusion one can draw from this is that all these are "rising" vowel complexes since the speaker seems to dwell on the second vowel while giving minimal time to the steady-state of the first. The brief articulation (and our perception of a short) vowel immediately succeeded by a different vowel of longer duration may be a marker of internal juncture.

The second important observation is the un-interrupted continuity of the formants of the first vowel into the corresponding formants of the second, with transitions of greater or lesser duration depending on the qualitative distance between the elements of the sequence in question. This is as it should be, as a situation whereby the formants of the first vowel did not connect with their counterparts in the second would be impossible.

The final point to make is that there seem to be no limits on the combinability of the five vowels of Central Shona. This suggests that we are dealing with sequences rather than unit phonemes such as diphthongs. In English, for example, there are far fewer combinations, which in itself suggests that the examples that do occur are unit phonemes. The twenty-nine examples described

above are not exhaustive of the possibilities in the language.

3.2.4. The grammatical function of the vowels in the sequences.

By "grammatical function" is meant the assignability of the vowel elements in a sequence to independent lexical items or other meaningful formatives. The first three examples described above:

11. [t`àé`ndà], "we have gone"

12. [t`àí`mbá], "we have sung"

13. [t`àó`ná], "we have seen"

are made up of elements assignable to different formatives as follows: the first two segments in each utterance [t-] and [-a] are the common denominators for the three examples. The [t-] stands for the person, namely the first person plural "we"; whereas [-a] indicates the tense of the verb in question, namely the immediate past "have". Therefore, taking the example 11 above, we can, illustratively, form paradigms using different "persons" in contrast with the first person plural in the subject position:

/t-à-é`ndà/ first person plural "we"

/nd-à-é`ndà/ first person singular "I"

/w-à-é`ndà/ second person singular "(s)he"

/v-à-é`ndà/ third person plural "they"

/m-à-é`ndà/ second person plural "you".

The vowel [-a-], the tense marker, can also enter into paradigms with other formatives marking different tenses.

For example the formative [-o-] can be placed in the same slot as that occupied by [-a-] above to mark the exclusive present indicative, translated "now I...":

/t-ó-éndà/, "now we go".

/nd-ó-éndà/, "now I go".

The second parts of the examples 11-13 above:

11. [-éndà]

12. [-ímba']

13. [-óná]

are the stems of the verbs to which the person-cum-tense markers have been added. This means that the second vowel of the sequences under discussion belongs to a different form, the stem of the verb. The two vowels therefore belong to different syllables.

The examples 22 through 25 can be analysed in much the same way as the examples above. The linking factor is that here the same person is used as in the examples just discussed, namely the first person plural, marked by [t-]:

22. [t-ó-índà]⁴³, "now we go"

23. [t-ó-érèrà], "now we float away"

24. [t-ó-à-tórá], "now we take them"

25. [t-ó-ù-tórá], "now we take it".

The first vowel of the sequence [-o-], marks the exclusive present indicative. In examples 22 and 23, [-índà] and [-érèrà], respectively, are the verb stems tonally inflected for tense. The lexical tone pattern is

/-índà/

/-érèrà/.

43. The forms /-éndà/ and /-índà/ are free variants in my speech for "go" (v.i.).

Examples 24 and 25 call for elaboration since they introduce a new element: their second vowels in the sequences do not belong to the stem of the verb in question. They are markers of object infixes.⁴⁴

In example 24 [t-ó-à-tórá], the second vowel of the sequence translates as "them" in the rendition "now we take them". In other words, it marks the noun class of the object of the verb [-tórá], "take". It indicates that the object is a noun belonging to class 6, just [-u-] in example 25 [t-ó-ù-tórá] tells us that the object of the verb in this case belongs to noun class 3. So both

24. [t-ó-à-tórá]

and 25. [t-ó-ù-tórá]

are "subject + (Tense) + object + verb" sentences of Karanga.

Our observations in examples 11 through 13, and 22 through 25, show that the following vowel sequences are made up of vowels belonging to different syllables, and the syllables belong to different grammatical formatives: -ae-, -a-, -ao-, -oe-, -oa-, -ou-.

The next set of vowel sequences to consider are those in the following examples:

19. [ndié], "it's him/her"

20. [ndíónè], "let me see"

21. [ndíumbè], "let me mould"

All these examples are made up of vowels which can enter into paradigms with different formatives resulting in meaning differences. Taking the second vowel of example

44. See Chapter 7 below.

19 first, this can enter into the following contrasts:

/ndi'-e/, "it's him/her

/ndi'-ni/ "it's me"

/ndi'-mi/ "it's you" (pl.)

/ndi'-we/ "it's you" (sing.)

/ndi'-tʃo/ "it's the one" (n.cl.7)

/ndi'-vo/ "they are the ones" (n.cl.8)

In this example, therefore, the second vowel stands for the third person singular in a deictic construction, and can enter into paradigms with other formatives in similar constructions but indicating different persons. And the first vowel is not affected. The structure [ndi'-] functions as one unit. Hannan (1974, p.440)⁴⁵ calls it a "copulative formative". The two vowels of [ndie] are, therefore, nuclei of two syllables abutting but belonging to different formatives.

The examples 20 and 21 above fall into a subclass of their own. The formative [ndi'-], when prefixed to the stems of verbs, refers to the first person singular "I" or "me". The forms [-'umbe]

and [-'one]

are derived from the stems of the verbs

/-umba/, "mould"

/-ona/, "see"

by rules of the phonological component whereby the final vowel of the stem /-a/ is changed to [-e]. The vowel change is syntactically conditioned and is made in order to express purpose or intention with the verb almost invariably preceded by /Kuti/, "so that":

45. Hannan, M. S.J. (1974) op.cit.

/ndìno'dá' Kùtì ndìpóre tsàmbà/, "I want to write a letter".

The form [ndì-]⁴⁶, indicator of the first person singular, (as opposed to [ndi-], "it's", the copulative formative) enters into paradigmatic relationships with other person markers as follows:

Kùtì	ndì-ónè	(stem-óná)	so that I may see
Kùtì	tì-ónè	"	so that we may see
Kùtì	mù-ónè	"	so that you may see
Kùtì	ʋá-ónè	"	so that they may see

Incidentally, there seems to be an interesting symmetry in vowel quality associated with the forms for

- (a) the first person singular on the one hand, and the first person plural on the other;
- (b) the second person singular, and the second person plural;
- (c) the third person singular, and the third person plural;

	<u>Singular</u>		<u>Plural</u>	
<u>First:</u>	ndì-ónè	"I"	tì-ónè	"we"
<u>Second:</u>	ù-ónè	"you"	mù-ónè	"you"
<u>Third:</u>	á-ónè	"he"	ʋá-ónè	"they"

For example, the vowel [i] is found in the form for the first person singular and the first person plural; and, similarly, the vowel [u], for the second person, and [a] for the third person, also singular and plural. This suggests that the language does not have a first person singular morpheme, and a first person plural morpheme as such; but rather an underlying first person morpheme, realised on the phono-syntactic surface structure

46. See Hannan, M. S.J. (1974) p.440 for the distinction.

as [ndi-] in the singular and [ti-] in the plural, and similarly for the second and third persons.

The examples just discussed, that is, 19, 20 and 21, show that the following vowel sequences are made up of elements belonging to different syllables and formatives: -io-, -iu- and -ie-.

There is another set of examples that are amenable to the same treatment as that applied to examples 20 and 21 above. They have a different set of vowel sequences:

26. [mu'-imbè] "you (pl.) may sing"

27. [mu'-à-torè] "You (pl.) may take them"

28. [mu'-éndè] "you (pl.) may go".

(These examples may also occur in command constructions: "you must.."). The common factors among the three examples are, first, that they all express purpose, and secondly, they all have a second person plural form prefixed to the verb stem. In examples 26 and 28, therefore, we have the first vowel associated with the second person form, [mu'-], "you" (pl.) and the second vowel is part of the verb stem.

Example 27 is slightly different in that the verb concerned is transitive. The form [-a-], coming between [mu'-], the subject of the clause, and [-torè] the verb, is, in fact, the indicator of the noun class of the object, namely class 6. The clause [mu'-à-torè] is a construction of the form:

Subject + object + verb,
the object slot being occupied by -a-, the pronominal and concordial form of such nouns as

/mà-pàngá/ "knives"

/mà-kumbò/ "legs"

This gives us another set of vowel sequences found in Shona speech within the same orthographic and phonetic form, but belonging to different syllables, which in turn belong to different grammatical formatives: -ui-, -ua-, -ue-.

Example 17, [sé-à-ngú], "like mine", is similar to example 27 [mú-à-tóré], "you may take them", in having [-a-] as the pronoun or grammatical concord. In a sentence such as /màkumbò ako akaita seangu/ "your feet are like mine", the /-a-/ refers back to /màkumbò/, "feet" and serves as the marker of grammatical concord, binding the syntagm together. If the addressee in the above sentence were to respond, he would most likely not use the noun /màkumbò/ but the formative /-a-/, as in

/á-ngù mákùrù/ "mine are bigger".

The first vowel of the sequence in /seangu/ functions as a unit with the preceding fricative consonant in /sé-/, "like". So the two vowels are nuclei of separate syllables belonging to separate formatives.

Example 29, [mù-òndé], "fig-tree", contains a sequence which can be separated as follows: the first vowel functions as a unit with the preceding nasal consonant in /mù-/: the noun prefix for class 3 nouns. As such, /mù-/ contrasts with the prefix of the nouns of class 4, the numerical plural of class 3 nouns:

/mù-òndé/ "fig-tree"

/mì-òndé/ "fig-trees"

The second vowel [-o-] is the initial segment of the stem.

The two vowels are, therefore, nuclei of separate syllables and formatives.

3.2.5. Sequences of vowels within a single lexical item.

The preceding remarks account for fifteen only of the vowel sequences under discussion. The remaining four

14. [táúrá] "speak"
15. [séi], "how", "why"
16. [sèúra] "weed!"
18. [ndiáni], "who is it?"

require a different approach. The first three examples (14, 15 and 16) contain sequences of vowels that cannot be separated by using morphological segmentation as in all preceding examples. The sequences are integral to the stems of the lexical items concerned. There is, however, syntactic or phono-syntactic evidence for treating the vowels in these utterances as nuclei of two independent syllables. This evidence is derived from the phenomenon of penultimate syllable lengthening mentioned earlier. In the following sentences:

- (a) /hándídí zókútáù:rà/, "I don't want to talk"
- (b) /záitá sé:i/, "what's the matter?"
- (c) /úsá séu:rà/, "don't weed!"

the second of the sequence of two vowels is lengthened. This means that it is the nucleus of the penultimate syllable. The other vowel in the sequence must therefore belong to a syllable of its own.

Example 18 [ndiáni] "who is it?", can be separated into two formatives [ndi-], the copulative

formative meaning "it is...", and [-àní], "who?"

The vowel sequences discussed in this chapter, therefore, fall into two broad categories: those made up of vowels which are assignable to the syllables of different formatives:

- | | |
|----------|----------|
| 11. -ae- | 12. -ai- |
| 13. -aə- | 17. -ea- |
| 18. -ia- | 19. -ie- |
| 20. -io- | 21. -iu- |
| 22. -oi- | 23. -oe- |
| 25. -ou- | 26. -ui- |
| 27. -ua- | 28. -ue- |

and those sequences which must be treated as forming the nuclei of two abutting syllables within the same lexical item or grammatical formative:

- | | |
|--------------|----------|
| 14. -au- | 15. -ei- |
| and 16. -eu- | |

This division into two categories is not watertight: it holds true only for the actual examples used in this discussion and others like them. But examples in the second category could be found in the first if a different set of utterances amenable to the methods used here to isolate examples of the first category were analysed, and vice versa.

In the first category we have two syllable nuclei corresponding to two different grammatical formatives: whereas in the second category we have two syllable nuclei corresponding to one grammatical formative or lexical item. In the former case we have both acoustic (formant transitions) and auditory and grammatical evidence for separating the vowels of each

sequence into two syllables belonging to two formatives; and in the latter case we have acoustic as well as auditory evidence (for the presence of two vowel qualities) for recognising two syllables, and grammatical or lexical reasons for assigning both syllables to one stem.

There is also the further evidence against positing unit phonemes here that the two vowels involved in each sequence make their own independent selection of tone.

Doke's (1931, pp.31-32, para 30)⁴⁷ treatment of the vowel sequences in Central Shona leaves much to be desired. He says "investigation has shown that at least two diphthongs *ou* and *ai* occur in certain Shona dialects". The nature of the investigation is not reported. But, as examples of words in which these so-called diphthongs occur, he quotes:

Zezuru:- *mhou*, "ostrich", "cow"

zou, "elephant"

and Zezuru and Manyika:- *mai* "mother"

jai "evil spirit"

hai "stop it" - actually "No!"

zai "egg"

In the first place, the examples *mhou*, *mai* and *zai* are also used in Karanga and Korekore. Manyika speakers use /*amai*/ for "mother".

Doke adds, "In each of these cases the diphthong when used, has a falling tone". This is true only for

47. Doke, C.M. (1931) op.cit.

some of the examples given by Doke, namely m^hóu, h^hai and z^hai. The word /nzòu/ has low level tone, so has /s^hai/; /m^hai/ has high level tone on each vowel.

In section 51, Doke gives examples of juxtaposed vowels, some of them with incorrect tone markings:

<u>Doke</u>		<u>Correction</u>
t ^h aurá	"speak"	-
b ^h gáirà	"wink"	b ^h gáirá
nánáirà	"toddle"	nánáirà
wáujà	"came"	wáujà "he has come"

(The correct translation for Doke's "he came" is wákáujá;
wáujà, Doke's example, translates "have you come?").

Kùtsáirá	"to sweep"	-
----------	------------	---

Of all the examples above, only the first and last have the correct tone marking.

However, the main point is that there is no difference in my pronunciation of the vowel complexes in

mái "mother" (HH) (Doke's "diphthong")	
and b ^h gáirá "wink"	} (his "juxtaposed vowels")
nánáirà "toddle"	

It is not clear on what basis he makes the distinction diphthong versus juxtaposed vowels.

3.2.6. The distribution of glides among vowel sequences.

Introduction. In the articulation of a sequence of speech segments, the organs of speech move very rapidly from the place of articulation of one segment to that of the next, normally by the most direct route. The distance between the places of articulation of any successive segments will determine the amount of time

spent (see the duration of formant transitions, Table II, pp.140-142 above) by the organs of speech in the process. If the movement of the organs is particularly rapid, their adjustment will not be audible; if slow, such movement will be audible. But whether or not it is audible, the "glide" exists since the organs move from one position to the next. Henry Sweet (1877)⁴⁸ says the glides are "implied" by the mere juxtaposition of any two sounds which are produced by different vocal tract configurations.

Thus in Karanga, the sequences of vowels within a form will either "imply", or produce audibly a glide depending on

- (a) the distance between the two successive positions,
- (b) the speed with which the movement is made.

Karanga speakers insert either a [j] or a [w] between the members of a vowel sequence. And since the syllable is always open in the language, the glide so inserted acts as the syllable onset of the second syllable. The "choice" between [j] and [w] (distinct from the semi-vowels of the next chapter) depends on the starting point of the tongue movement towards the next vowel quality, relative to two parameters: backness and height.

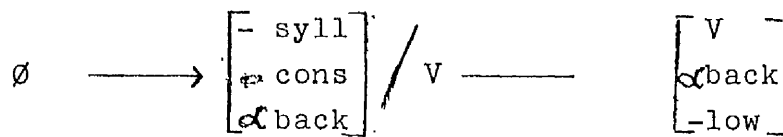
The glides in the following pairs of vowels are weak or inaudible because the distance between the vowels is short or there is no movement at all:

- (a) slight movement: -ei-, -ie- → [j]; -ou-, -uo- → [w]
- (b) no movement: -aa-, -ee-, -ii-, -oo-, -uu-.

48. Sweet, H. (1877) A Handbook of Phonetics, Excerpt in E.J.A. Henderson (ed.) *op.cit.*, p.152.

The rule for examples in (a) is

Rule 1



1 2 3
Condition 1 \neq 3.

The sequences in (b) were not analysed above. But the members of these sequences can be assigned to different syllables by the methods used for separating the other examples.

3.2.7. (a) Examples that allow an optional [j] glide.

An optional [j] off-glide for the first, and [j] on-glide for the second vowel may be inserted between the members of the following sequences:

-oe-, -oi-, -ue-, -ae-, -ai-, -ui-.

In all the cases, the movement is towards a front unrounded mid or high tongue position.

(b) Examples that allow an optional [^w] glide.

There may be a [^w] off-glide for the first vowel and a [^w] on-glide for the second vowel in the following sequences:

-eo-, -eu-, -io-, -iu-, -uo-, -ao-, -au- and -ou-.

In all these cases, the movement is towards a back, mid or high tongue position accompanied by lip rounding.

The fact that the off-glide from the first vowel, which is the on-glide of the second vowel in a sequence

(in acoustic terms, the formant transition) has a quality very close to that of the second vowel, persuaded me, in section (a) above, Table II, to add the duration of the transition between the vowels to the duration of the steady-state of the second vowel.

3.2.8. Examples allowing the insertion of neither glide.

Among the examples discussed in this chapter I found three vowel sequences between which neither [ʝ] nor [ʷ] can be inserted: -ea-, -oa-, -ua-. But I have to make one qualification: -ea-, apparently, does not admit a glide only when the -a-, as in [sé-à-ŋgú], "like mine", is the grammatical concord or pronominal form of class 6 nouns and not part of the stem of a lexical item. Otherwise it is permissible to insert a [ʝ], as in the word borrowed from English:

/tʃéà/ and /tʃéʝà/, "chair", being free variants in Karanga.

The common denominator among examples that do not admit either glide is that they all involve movement towards the open non-back, low tongue position, predominantly from the back, rounded, mid or high tongue position.

In view of the /-ea-/ exception just mentioned, the rule for glide insertion will be rewritten as:

Rule 2

$$\emptyset \longrightarrow \begin{bmatrix} \text{syll} \\ \text{cons} \\ \alpha_{\text{back}} \end{bmatrix} \quad / \quad \underset{1}{V} \quad \underset{2}{\text{---}} \quad \underset{3}{\begin{bmatrix} V \\ \alpha_{\text{back}} \end{bmatrix}}$$

Condition 1 ≠ 3.

Whether they are audible or only slightly so, glides in these positions are not linguistically, or even socially, significant. Thus

$$\left. \begin{array}{l} /ndóit\grave{a} \acute{s}éi/ \\ \text{and } /ndó^{j\acute{}}it\grave{a} \acute{s}éj\grave{i}/ \end{array} \right\} = \text{"what shall I do?"}$$

In other words, [^j] here cannot be equated with /t/ in the same utterance. If the latter is replaced by /^g/, for example, the meaning is changed radically:

/ndóig\grave{a} \acute{s}éi/, "how shall I eat it."

But replace [j] by any consonant in the Karanga phoneme inventory and the result is a nonsense form: conversely, remove the glide and leave a 'blank' and the meaning is preserved. Its use has no social consequence either, since the speaker who uses it or omits it is not a more or less acceptable member of any group solely by that token. On the other hand, my [^htáw^hn] as opposed to RP [^htáw^hn], "town" is an index of my Shona background.

3.3. Conclusion

Diphthongs and juxtaposed vowels in Karanga.

Two types of argument can be adduced for the case against the positing of diphthongal phonemes in Central Shona: one phonological and the other syntactic.

3.3.1. The phonological argument.

As already noted above, Doke reports that his investigations established that "there are at least two diphthongs in Shona *ou* and *ai*". This kind of interpretation would mean that the examples quoted by him, namely

mái "mother", zái, "egg"

nzòu "elephant", etc. are monosyllabic words of the structure CV.

My first argument against this interpretation derives from what I have said above concerning the distribution of the glides [j,^w] among the vowel sequences, whether or not such sequences belong to one lexical item or to two formatives bound together in the spelling but having distinct functions. The [j] glide can be inserted between the vowels of

mái ɤái zái

to give the equivalent forms

ma^ji, ɤa^ji za^ji.

Similarly, it has been shown above that the vowel sequence "ou" belongs to a class of sequences between which the glide [^w] can be inserted thus

m^hóu ~ m^ho^wu
nzòu ~ nzò^wu, and so on.

This means that these words are not monosyllables: if they were, the option of inserting a consonantal glide between the elements of the so-called diphthongs would not be there.

3.3.2. The Syntactic Argument.

This argument also derives from what has been said in a previous section: that in Central Shona, the vowel forming the nucleus of the penultimate syllable in an utterance (I find this particularly true of declarative statements) is longer in duration than all the other vowels in the same utterance. This fact of

Central Shona phono-syntax has important implications for the problem under discussion.

Penultimate vowel lengthening means that in the sentence /ndàwónà ʃumbà/, "I have seen a lion" the vowel /u/ is marked for length:- /u:/. This vowel is the nucleus of a penultimate syllable since there is no doubt about /ʃumbà/, "lion", being disyllabic.

If, however, the word /ʃumbà/ were replaced by a monosyllabic word, the penultimate position would then be occupied by the last syllable of

/ndàwónà/, "I have seen":

e.g. /ndàwónà: mbyá/, "I have seen a dog".

Doke's interpretation treats /máí/, ʃáì, and zái/ as monosyllables with diphthongal nuclei. This would mean that if we place any one of these words in the same slot occupied by /mbyá/ in the frame

/ndàwónà ——— /, the syllable /-na/ of /ndàwónà/ would retain its prominent length since it would still be in a penultimate position in the utterance. But this does not happen. In the sentence

/ndàwónà zái/, "I have seen an egg",

penultimate lengthening passes on to the first vowel of /zái/, "egg". This means that in this statement, /-na/ is no longer the penultimate syllable: the syllable /za-/ is penultimate, and /-i/ is the nucleus of a separate, in this case, the final, syllable. So /-ai-/ is not a diphthong, nor is /-ou-/: there are no diphthongs in Central Shona.

The frame used in this argument may be shown as follows - the tones marked are those of the nouns in isolation.

<u>Subject-verb slot</u>	<u>object slot</u>	
(a) ndàwóná'	{ ʃúmBà ŋg ^w é:nà nú:nzi }	"lion" "crocodile" "fly"
(b) ndàwóná'	{ ŋg ^w é mbá tso yú nda pxá ná }	"leopard" "house" "kidney" "soil" "louse" "sugar cane" "foot crack"
(c) ndàwóná'	{ má:i zá:i ¹ ʃá:i nzo:ù ¹ mó:ù ¹ }	"mother" "egg" "evil spirit" "elephant" "cow"

(1. In cases where the releasing consonant is breathy-voiced, it is only the first vowel of a sequence that shares some of the breathiness: both would be affected if they belonged to the same syllable.)

Observations.

(i) In (a) above, the vowel lengthened is that in the first syllable of the disyllabic object noun: this is the penultimate syllable.

(ii) In (b), the last syllable of /ndàwóná/ is in the penultimate position since the object noun, the last

word of each sentence, is monosyllabic: therefore /-na/ is lengthened to /-na:/.

(iii) In (c), the syllable /-na/ is no longer in the penultimate position and therefore loses its extra length of the examples (b). The first vowel of the object item is lengthened, thereby behaving like that of the disyllabic object items in (a). The object items in (c) are, therefore, disyllabic.

CHAPTER 4
THE SEMI-VOWELS

4.0 Introduction.

The term "semi-vowel" is used to refer to the phonemes of Karanga symbolised as /w, w̥, ʋ, j/. Throughout I shall reserve the term "glide" for the adventitious sounds produced by the organs of speech when moving from their "posture" for one speech sound to that for another. As I have said, in the previous chapter, such glides were discussed extensively by Sweet (1877)¹,

"It would clearly be impossible to symbolise all the infinitesimal intermediate positions of which a glide is made up, nor is it ever necessary, the general principle being that in all cases of transition from one fixed position to another, the shortest way is taken".

And,

"Glides are sounds produced during the transition from one sound to another. Thus in kii, "key", we have the glide from the K-position to the ii-position, which does not, however, require to be written as it is implied by the positions of K and ii."

Semi-vowels, considered as consonantal phonemes, are essentially radically different from these incidental phonetic phenomena described by Sweet. Here I am not discussing those glides which it is optional for a Karanga

1. Sweet, H. (1877) op.cit.

(1891) op.cit., Excerpts in E.J.A. Henderson (1971) op.cit., pp.60-85,229, respectively.

speaker to insert between juxtaposed vowels described in the previous chapter. Although semi-vowels are themselves characteristically much shorter in duration than many other speech sounds (Gimson (1962 p.212.)² says "A semi-vowel is a rapid vocalic glide on to a syllabic sound of greater steady duration.") the most important difference between them and glides is of duration: semi-vowels having longer duration than these glides or transitions. Note that Gimson uses the term "glide" to refer to the rapid movement of the organs of speech, and not to the sounds so produced. It is necessary to draw this distinction between these two types of sound since some writers use the terms "glide" and "semi-vowel" interchangeably.

4.1. The Labio-velar semi-vowels /w/ and /ɰ/

Pronunciation:

The first semi-vowel of Karanga, the labio-velar /w/, is produced by raising the back and root of the tongue towards the soft palate and by simultaneously rounding the lips closely. Air from the lungs passes through the glottis, causing the vocal folds to vibrate. There being a velic closure to cut off nasal resonance, this periodically vibrating air flows into the oral cavity and out of the speech tract through the mouth, using a median mode of egress.

Generally speaking, the velar stricture and the

2. Gimson, A.C. (1962) op.cit.

lip-rounding are narrower than those for other vocalic sounds produced in these places of articulation. For example the tongue raising towards the soft palate is slightly more radical than for the high back vowel /u/; and the lip-rounding is closer than for the same vowel. But the velar stricture for /w/ is not as narrow as that for the velar fricatives /x, ɣ/, or, for that matter, the velar plosives /k, g/.

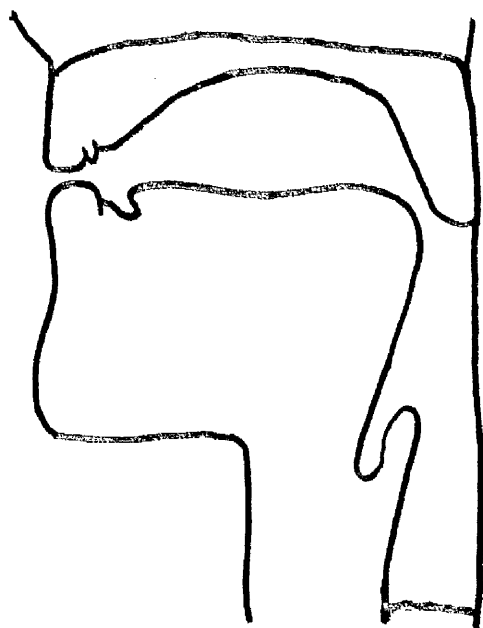


Diagram 6. /w, ʋ/

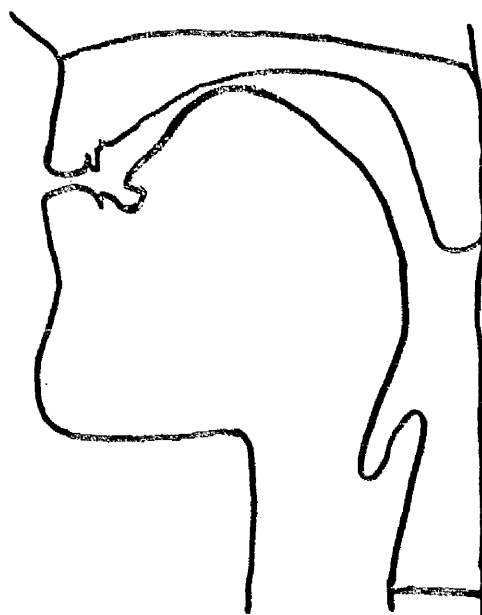


Diagram 7. /j/

The tip of the tongue remains in a position of rest behind the lower incisors (see Diagram 6).

4.2. The Palatal semi-vowel /j/

For the palatal semi-vowel /j/, the glottal and velic strictures are the same as for the labio-velar semi-

vowel /w/. The only difference is that the larynx is raised to a higher level since the front and blade of the tongue are raised towards the hard palate. This raising is to a closer degree than for the front unrounded vowel /i/. The lips are spread and the tip of the tongue remains in a position of rest behind the lower incisors. This stricture is also maintained only for a short time (see Diagram 7).

The foregoing might seem to suggest that there is a fixed degree of proximity (between the back of the tongue and the velum for /w/, and between the front of the tongue and the hard palate for /j/) from which the organs of speech operate. But, as Gimson (1962, p.212)³ says, "The actual point at which the essential glide begins depends on the nature of the following sound." In English, for example, "the glide of /j/ to /i:/ in 'yeast' has a closer beginning than that of /j/ to /ʊ/ in 'yacht'..." In other words, the articulations assimilate somewhat to the 'place' of articulation of the following sound.

Although from a purely phonetic point of view, these sounds are generally vocalic, that is, they have no consonantal constriction and have a vowel-type formant structure, they are phonologically regarded as consonantal phonemes because of their function in the syllable: they are marginal rather than nuclear to the syllable. And on average, vowels have longer duration than semi-vowels.

3. Gimson, A.C. (1962) op.cit.

Distribution:

The following are some of the lexical items in which these phonemes occur. There is no restriction on the sequence of either semi-vowel with the vowels of the language.

<u>Word initially</u>	<u>Intervocalically</u>
/w/:(i) /-wònèkà/, "bid farewell"	(i) /màwàrà/, "temerity"
(ii) /-wóná/, "see"	(ii) /màwéré/, "rapids"
(iii) /-wànà/, "find" or "get married to"	(iii) /kùwàjàwàjà/, "kill time"
(iv) /-wòmbèkà/, "baptise"	(iv) /màkùwà/, "rumours" or "gossip"
(v) /-wìrìràà/, "get on well"	(v) /màkùwèrèrè/, "outstanding spectacle."
/j/:(i) /-jàmùrà/, "help"(v.t)	(i) /-sìjána/, "part company with"
(ii) /-jàníká/, "leave on sun to dry"	(ii) /-féjáféjà/, "investigate secretly"
(iii) /-jàmbùkà/, "cross river"	(iii) /m ^w èjá/, "air" or evil spirit"
(iv) /-jèmùrà/, "show respect towards"	(iv) /mojò/, "heart" or "greed".

The fourth example of the occurrence for /w/ in an intervocalic position shows that there is a breathy-voiced labio-velar semi-vowel in Karanga. The breathy-voice is obviously not an environmentally conditioned feature since both /w/ and /w̥/ occur before the same vowels. The two phonemes are in contrastive distribution, as the following minimal pair demonstrates:

/kùwá/, "to fall"

vs.

/kùw̥á/, "to hear"

Other occurrences of /w/ are as follows:

- | | |
|-----------------------------|---|
| (i) /wí/, "voice" (n.5) | (i) /gùwà/, "rumour" |
| (ii) /wemà/, "scent" | (ii) /kùwífurá/, "to snatch" |
| (iii) /wénà/, "edible root" | (iii) /nùwà/, "to smell" (v.i.) |
| (iv) /wámándá/, "trumpet" | (iv) /màwék ^w é/, "chance meeting" |

4.3.0. Acoustic Studies of Semi-vowels

P. Delattre (1968)⁴ studied French consonants with a view to finding a set of distinctive features based on acoustic cues. He lists a number of acoustic features which characterise semi-vowels, using the term "glide" for "semi-vowel".

The French phoneme /w/, along with other phonemes with a labial stricture, has an F2 transition with a locus at about 700 Hz. The defining difference between /w/ and other labials is that whereas the formant transitions for these other consonants merely point towards their loci, those for /w/ actually touch their loci.

The French palatal semi-vowel /j/ has an F2 frequency transition with a locus at 3 KHz. And the F3 frequency transition has a locus at about 3.5 KHz. Adding that /j/ is the dorso-palatal consonant par excellence, Delattre says that this phoneme shares the F2 and F3 transition loci with /p/, "though they have different modes of articulation." And, finally, "The links of the glides are continuous, i.e. they

4. Delattre, P. (1968) "From Acoustic Cues to Distinctive Features". Phonetica 18, pp.198-230.

are in continuity with the formant transitions which they join together without a break." This is reminiscent of the formants of the juxtaposed vowels of Karanga discussed in the last chapter.

The results of research reported by O'Connor et al. (1957)⁵ are of particular interest because synthetic sounds were used in order to isolate those acoustic features which demarcate semi-vowels from other sounds and from one another. Since such sounds are not, strictly speaking, inextricably associated with a particular language, they can be safely used as objective reference points, like cardinal vowels, in the description of the sounds of particular languages. The authors used synthetic speech because

"This provides a convenient method for making experimental modifications in various presumably important parts of the pattern and then evaluating the effects of these changes on the sound heard."

They worked exclusively with nonsense syllables of the pattern CV. And they used seven vowels "corresponding approximately to the cardinal vowels [i, e, ɛ, a, ɔ, u]".

Their paper is divided into two parts. Part I consists of a summary of the results of the experiments with synthetic speech. The authors themselves evaluate the effect on the stimuli (the sounds perceived) of changing certain acoustic features. In Part II, forty-four naive listeners were asked to identify the sounds

5. O'Connor, J.D., Gerstman, L.J., Liberman, A.M. Delattre, P. and Cooper, F.S. (1957) "Acoustic Cues for the Perception of Initial /w, j, r, l/ in English" Word, 13, pp.24-43.

produced by the PPB (Pattern Play Back) as /w,j,r, or l/. This latter exercise was designed to test the conclusions which the authors had arrived at in Part I.

They first paid attention to formant transitions, and found that the second formant transition could be used to distinguish /w/: /r,l/: /j/. For /w/, the transition originates at a low frequency; for /r,l/ it starts near the middle of the frequency range 0-3 KHz, while /j/ requires that the transition start even higher. In addition, it was found that at the beginning of each formant a relatively short steady-state was required and they called it "the steady-state onset". These Onsets represent consonant loci, and are the acoustic counterparts of the starting points of the consonant articulations. "It is of considerable interest that the loci are explicit for /w,j,r,l/ because in this respect these four phonemes are different as a class from the stop and nasal consonants." For these other consonant classes, the authors found that the second formant transition does not begin at the locus but at a point somewhat delayed in time. The formant transitions for /w,j,r,l/ not only begin at the locus but spend from 30 to 50 msec there before proceeding to the steady-state positions of the vowels.

4.3.1. The number of Formants Needed (see Fig. 6, p. 183)

The difference between /w,j/ on the one hand, and /r,l/ on the other is that whereas the former can be synthesised satisfactorily with only two formants, the latter require three. In the case of /w,j/, this is to

be expected, bearing in mind the well-known possibility (see Chapter 3) of two formant vowel synthesis and the close articulatory and, therefore, acoustic resemblance between /i/ and /j/, and between /u/ and /w/.

4.3.2. Duration of Steady-state Onsets.

When O'Connor et al, omitted the onset and began the transition immediately, the resulting stimulus appeared to have an explosive beginning. They found that a 30 msec duration for the onset gave ideal results for /w,j/.

4.3.3. The starting frequencies of the transitions (steady-state onset frequencies).

(i) First formant onset. This is not distinctive among /w,j,r,l/ since, for the same following vowel, its form is unchanged throughout. The first formant onset for /w,j/ is in the neighbourhood of 240Hz, that is, near the first formant frequency for both [u] and [i].

(ii) Second Formant Onset. Omibus paribus, high frequencies in the region of 2.75 KHz give /j/, and low frequencies in the region of 600 Hz give /w/. The intermediate range gives either /r/ or /l/.

4.3.4. Transition Duration

It was found that the duration of the transitions does not differentiate these sounds among themselves, but distinguishes them from all other sounds. For example, very short transitions caused confusion with nasals and

stops, while with transitions which occupy too much time in moving from the steady-state onsets to the steady-state vowel formants there was the danger of losing the consonant impression in favour of a vowel of changing quality. A transition duration of 100msec was found to be suitable for the F2 and F3 transitions of all the four phonemes.

The paper just summarised has a lot of relevance to the study of the semi-vowels of Karanga. Whether the latter are similar to, or differ from, the sounds investigated by those scholars, the comparison will throw light on the phonemes I am describing. For example, how do the centre frequencies of the F1 and F2 of /j/ differ from those of the corresponding formants of /i/, and those of /w/ from those of /u/? Secondly, that the lower limit of the F2 onset for a synthetic and recognizable /j/ should have the same frequency, 2.76KHz, as the second formant of [i], as reported in that paper, means, in fact, that the F2 of /j/ is generally higher in frequency than the corresponding formant of the vowel.

If this is found to be the case with the natural-speech sounds of Karanga, the following would be a plausible physiological explanation: there is a narrower stricture between the front of the tongue and the hard palate during the articulation of /j/ than during that of /i/. This difference is obviously the cause of the different size and shape of the supra-glottal cavities for the respective sounds, leading to distinct resonance characteristics.

In the following analysis of the spectrograms of the semi-vowels of Karanga, a number of features will be specified: the formant structure of each semi-vowel, e.g. the number and centre frequencies of such formants; the duration of the steady-state onsets (if they are visible) and the duration of the vowel formant transitions.

4.4.0. Acoustic Features of Karanga Semi-Vowels

Research for this section consisted in selecting a number of lexical items and grammatical constructions in which the semi-vowels /w/ and /j/ are flanked by allophones of the same vowel phoneme. In all the examples chosen the semi-vowels are in an intervocalic position: VCV. The examples selected were recorded,⁶ pronounced by me in controlled laboratory conditions.

For /j/, the following examples were used (the number against each example is the number of the spectrogram):-

30. [tʃíjì], "what is it?"
31. [ájám̀b̀ùk̀à], "(s)he has crossed (the river)"
32. [mòjò], "heart" or "greed" for food
33. [m̀ùùjù], "baobab tree"
34. [séjèdù], "like ours" cl.9.

6. Examples used throughout this thesis were recorded in an anechoic room. After each recording session they were played back so that a judgement could be made as to whether they were an acceptable sample of Karanga speech. Otherwise the examples were re-recorded.

These utterances were analysed using the sound spectrograph. The spectrogram of utterance 30. shows that [j] has a clear formant structure, three formants being clearly visible: F1 at 250Hz, F2 at 2.30KHz, and F3 at 3.40KHz. Compare these frequencies with the centre frequencies of the vowel [i] on Table I(a)p. 134 : whose F1 is at 225Hz; F2 at 2.50KHz and F3 at 3.50KHz. On spectrogram 30, the F2 of the second vowel [i] in fact falls fairly steeply to reach its own steady-state from the corresponding formant of the semi-vowel.

One point that emerges from inspecting this spectrogram is that it is easier to make the minute demarcations on the spectrum reported by O'Connor et al. when working with synthetic, manipulable speech, than when one is analysing the spectrograms of natural speech. For example, the steady-state onset, which those authors found to be so crucial for a successful synthesis of the semi-vowel phonemes, cannot be detected on spectrogram 30. However, since I am at this point dealing with [j] in the environment of the front, high, spread vowel [i], failure to identify this feature is perhaps understandable: the onset is, after all, the acoustic counterpart of the starting point of the articulatory movement. The distance between [j] and [i] may not be long enough to produce an onset of a duration comparable to what O'Connor et al. found to be essential.

The duration of the steady-state of [j] seems to be about half that of the preceding vowel: 15 csecs and 30 csecs respectively. And the second and third formants of the vowels flanking [j] have a darker outline

presumably due to their greater intensity. On Spectrogram 31 [ájámbùkà] "(s)he has crossed (the river)", the semi-vowel [j] is surrounded by two realizations of the vowel phoneme /a/. The difference in the vocal tract configurations for [a] and [j] is reflected in the clear and steep formant transitions between the two sounds. On all these spectrograms the formant transitions of the vowels surrounding [j] have high explicit loci, so that the overall spectrum of $V_1j V_1$ has the shape of Figure 7:

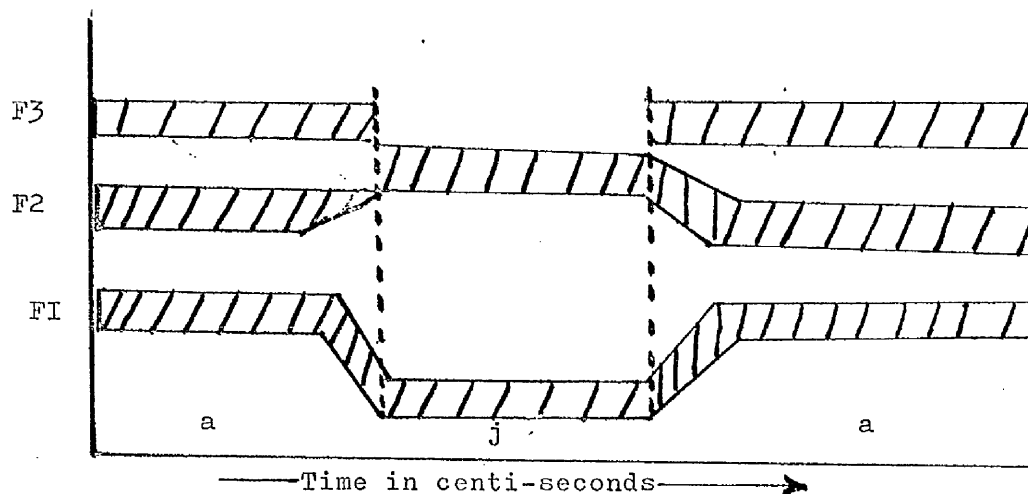


Figure 7. The formant transitions from $[V_1-]$ to $[-j]$ and from $[j-]$ to $[-V_1]$.

As we have seen in the chapter on vowels,⁷ the first two formants of the vowel [a] have centre frequencies at 725Hz and 1.50KHz respectively. The first two formants of [j] on spectrogram 31, on the other hand, are much

7. The vowels used in chapter 3 were pronounced in isolation, whereas the vowels surrounding the semi-vowels in this chapter are part of whole utterances. And the two sets of examples were recorded at different times during the research for this thesis. For these two reasons, there are likely to be slight differences in the spectra of the vowels in spectrograms 1 through 10 and those of the vowels in spectrograms 30 through 46: this is allophony.

farther apart than this: at 250 Hz and 2.5KHz respectively. So whereas the F1 of [a] falls in frequency to reach the steady-state of the corresponding formant of [j], its F2 rises in frequency to reach the steady-state of the F2 of [j].

The duration of the formant transitions between the vowels and the palatal semi-vowel in all cases averages 6 centi-seconds. The vowel has greater steady-state duration than the semi-vowel: 12 centi-seconds and 10 centi-seconds respectively on spectrogram 31. The observations made with respect to spectrograms 30 and 31 also apply to spectrograms 32 through 34, the most consistent feature being that the vowel formant transitions touch their loci in all cases, at 2.5KHz for F2, and 3.5KHz for F3.

The following utterances were recorded and analysed using the sound spectrograph in order to find acoustic features which characterise the labio-velar semi-vowels /w, w̥/:

35. [i^hw̥i^hsé], "make it fall!" (cl.9).
 36. [sèwé], "like you". 37. [m̩aw̩àrà], "temerity"
 38. [m̩úw̩újè], "you(pl.) ought to come".
 39. [m̩ów̩ónèkà], "now you(pl.) bid farewell."
 40. [k̩ùw̩á], "to fall". 41. [k̩ùw̩á], "to hear".

Acoustic features observed on the spectrograms of utterances with [w], and their difference from those of [j], are related to the physiological activity which produces these two sounds. The first obvious feature distinguishing these two sounds is the centre frequency of the second and third formants. These formants have much lower centre

frequencies at 1.00KHz and 2.50KHz, respectively for [w], than for [j] above. It would seem that the smaller oral cavity in front of the stricture for [j] is more conducive to the production of high frequency F2 (and F3) sounds than the larger oral cavity for [w] (although the latter cavity has a smaller aperture).

The first formant of [w] is, however, almost as low in frequency as that of [j]. This being the case, there is no visible formant transition between [i] (the vowel related acoustically to [j]) and [w], for this formant. The formant transitions of F2 and F3 of [i] both before and after [w], on the other hand, are quite steep on spectrogram 35. Both F2 and F3 of [i] before and after [w] in [i^ˈwɪs^ə] fall from their steady-state frequency to touch their loci at 1.00KHz and 2.5KHz respectively. These transitions take about 9csecs to accomplish. The steady-state of [w], (F2 and F3), is much shorter in duration than that of the surrounding vowels, and more especially the second vowel.

Spectrogram 36 [s^əw^é] shows much the same features as, but with bolder outlines than, spectrogram 35 above:

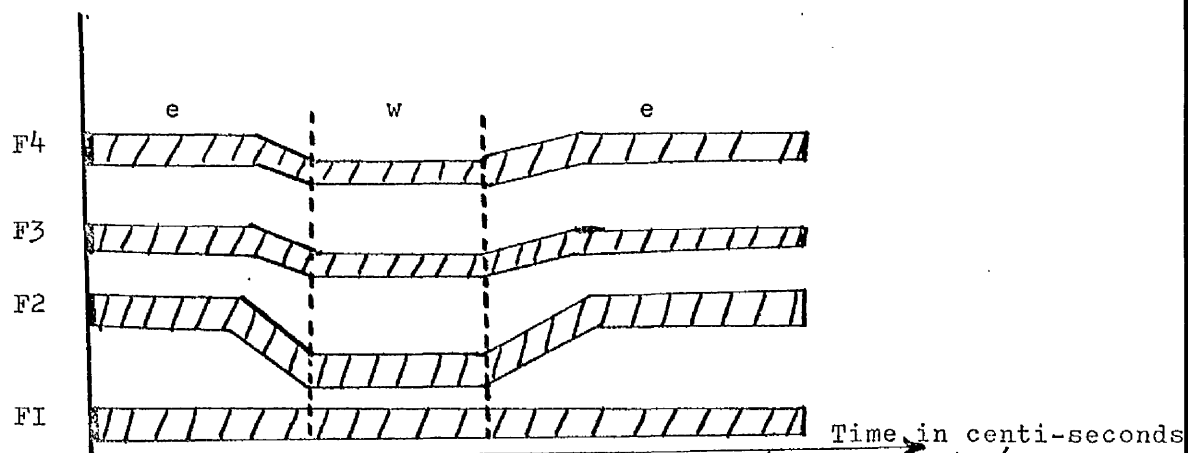


Figure 8. The Formant Structure of [-əwé] as in [səwé] "like you"(sing.), Spectrogram 36.

The second difference between spectrograms 35 and 36 is that the formant transitions between the adjacent vowels and [w] are less steep and less rapid on spectrogram 36. But both F2 and F3 of [e] fall to touch their respective loci for [w] when preceding it, or rise from these loci to reach the steady-state of the vowel coming after, this semi-vowel. The transitions of the preceding vowel formants seem to take longer to accomplish than those of the following: their durations being 9 csecs and 6 csecs respectively.

On spectrogram 37 [màwàrà], where [w] is surrounded by allophones of /a/, the vowel formant transitions are even less steep and take less time to accomplish than in the preceding examples. This, no doubt, has something to do with the tongue positions for [w] and [a] not being so radically different. In this example the first formant of the vowel falls in frequency before [w] and rises after it, the vowel's steady-state having a higher centre frequency than the corresponding formant of the semi-vowel.

On spectrograms 38 and 39, [múwújè] and [mówónèkà] respectively, where [w] is surrounded by back, lip-rounded vowels, there are no clear formant demarcations between the vowels and the semi-vowel because there are no transitions. Spectrogram 39 is interesting mainly because of the terminal transitions of F2 and F3 of [e] before, and of the initial F2 and F3 transitions of [a] after, the voiceless velar plosive [k]. These transitions point towards a common locus at about 2.5KHz, a feature which

has been found to distinguish velar consonantal articulations (see chapter 5, pp.198-333).

Spectrograms 40 [kùwá] and 41. [kùwá] were obtained in order to study the acoustic differences between the plain-voiced semi-vowel [w] and its breathy-voiced counterpart [w̥]. The one outstanding difference between these two types of labio-velar sound is that the striations of [w] are closer together than those of [w̥]; this relates to the fact that, auditorily, the high tone on the second syllable of example 41 is realised as a rising pitch, whereas that of the corresponding syllable on spectrogram 40 is high and level.

The difference in the distance between their striations, and the effect this has on the pitch of the syllable which they release, are the principal phonetic and phonological features distinguishing plain-voiced from breathy-voiced consonants - as I shall have occasion to elaborate in the following chapters.

All Shona words beginning with this breathy voiced semi-vowel are entered separately in Hannan's (1974)⁸ Standard Shona Dictionary, where they are spelt "hw", while those beginning with plain-voiced /w/ are entered under "w-". This separation is both phonetically and phonologically acceptable to me.

4.4.1. The Labio-dental Frictionless continuant, /w̥/.

Pronunciation. The labio-dental stricture for the pronunciation of this sound is very similar to that for

8. Hannan, M. S.J. (1974) op.cit.

/f,v/, the voiceless and breathy-voiced labio-dental fricatives. The edge of the upper teeth, the incisors, articulates with the inside surface of the lower lip. The main difference is that whereas this stricture is narrow enough to cause lung air to issue with friction for /f,v/, it is wide enough for the same volume of air to escape without friction for /ʋ/.

This latter sound does not have a voiceless or breathy-voiced counterpart. Throughout its pronunciation, the soft palate is in a raised position, thereby cutting off the nasal chamber, and the vocal cords vibrate periodically, producing voice. The body of the tongue assumes its position for the adjacent, especially the following, vowel. The tip of the tongue remains in a position of rest behind the lower teeth.

Distribution

Like all the other consonantal phonemes in Karanga, the labio-dental frictionless continuant occurs in CV-syllables, where V stands for any one of the five vowel phonemes, /i,e,a,o,u/. The difference between this consonant and most other consonants in the language is that unlike the rest, it does not occur in consonant clusters.

This consonant is, in the speech of most people in the Karanga area in free variation with the bilabial frictionless continuant /β/.⁹ Thus, in the following

9. /β/ is the symbol for the bilabial fricative with voice. I am using the subscript hook to indicate a stricture of wide approximation, that is, frictionlessness.

contexts /v/ may be replaced by /β/ with no semantic changes:

- | | |
|-----------------------|-------------------------------|
| /-vòrà/, "rot"(v.i.); | /-vává/, "taste bitter"; |
| /-vìrà/, "boil"; | /-vèrèngà/, "read" or "count" |
| /tʃìvúnó/, "waist" | /mávókó/, "hands" or "arms" |

Acoustic Features

A number of utterances containing the labio-dental frictionless continuant were recorded and analysed using the sound spectrograph:

42. [mùvìrì], "body"(n.cl.3);
43. [sévenzá], "work" (v.i.).
44. [mávátumà], "you(pl.) have sent them"
45. [mùvùjú], "baobab tree".
46. [sóvópéngá], "as if they were insane".

Among the acoustic features observed on the spectrograms of these utterances, there are three which link the labio-dental sound with the semi-vowels /w, ɰ, j/ discussed in the preceding sections, hence its inclusion in a chapter on these sounds. These three features are, first, the clear vowel-type formant structure of the sound; secondly, the continuity, with some transitions, of these formants without a break into the formants of adjacent vowels and, finally, the fact that the transitions of the formants of such vowels do not merely point towards, but actually touch, their respective loci: the transitions have "explicit" loci. The last feature, as we have observed above, marks off the semi-vowels as a class apart from all the other consonants.

The formant structure of the semi-vowel [v] is clearly displayed on all the spectrograms 42 through 46.

On spectrogram 44, for example, three formants are clearly visible: F1 at 400 Hz; F2 at 1.25 KHz and F3 at 2.40 KHz. There are other higher frequency formants on this spectrogram but they are weak in intensity.

The formants of the vowels surrounding [ʊ] have transitions touching their loci at the centre frequencies of the three formants of [ʊ] above. The fact that the F2 and F3 transitions of the vowels have loci at 1.25 KHz and 2.40 KHz is rather confusing since, [ʊ] being a labio-dental articulation, I had expected these loci to be lower in frequency.

Figure 9 is a diagrammatic representation of the formants of [-ʊir-], the middle portion of spectrogram 42:

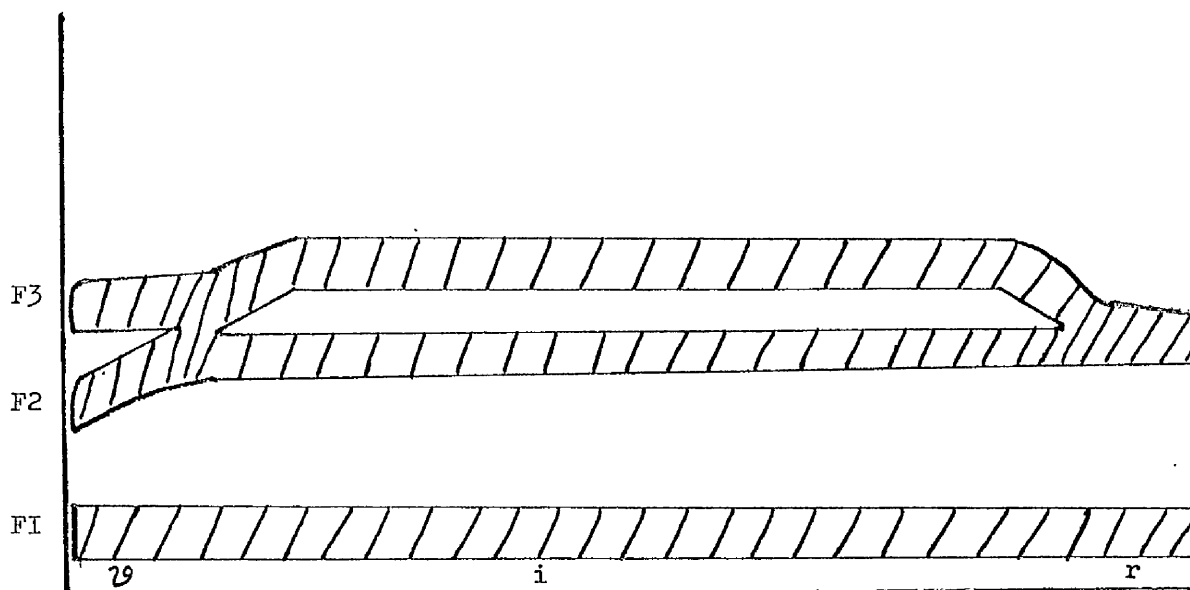


Fig. 9. [-ʊir-] of [mùʊìrì]—Time in centi-seconds—→

On that spectrogram, the initial F2 and F3 transitions of the vowel [i] bend downwards to touch their loci for [ʊ]. The second formant of the latter segment does not have a long steady-state duration, but bends upwards

to meet the corresponding formant of [i] at about 1.5 KHz.

All three acoustic features listed above are clearly visible on spectrogram 43. [-sévéŋzá]. The second and third formant transitions of the vowel [e] both before and after [ʊ], touch their loci at 1.55 KHz and 2.50 KHz respectively, although I had expected them to be much lower, as in the case of [w], which also has a labial stricture. When Fig. 10 below is compared with that of [w] in the context [e-e] (Fig. 8) it is clear that the loci of F2 and F3 transitions are higher in frequency for [ʊ]. This may be because in the latter case the tongue maintains its position for [e] during the pronunciation of [ʊ], whereas it has to move from the [e] position to that for [w] and back again in pronouncing [sévé].

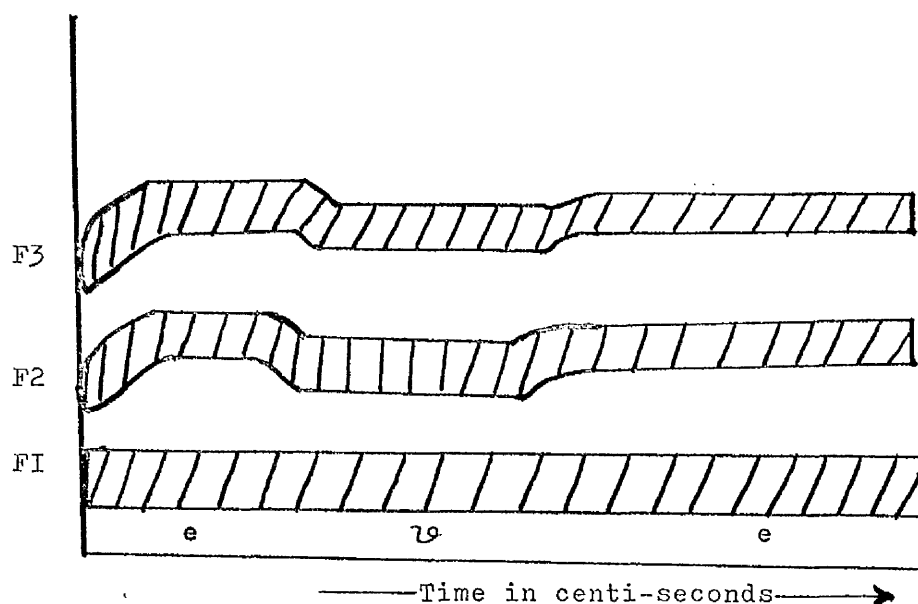


Fig. 10. Showing the first three formants [-évé] as on Spectrogram 43. [sévéŋzá].

I must underline one fact though: that the frequencies of these transitional loci meet expectations in that they lie between those for [w] on the one hand, and those for [j] on the other. This corresponds to the labio-dental place of articulation being between the bilabial([w]) and the palatal ([j]). What seems a little strange is that the loci should be closer to those for [j] than those for [w]. This leaves the fact that the vowel formant transitions touch their loci as the major acoustic feature distinguishing [ʋ] from the alveolar nasal [n], for example.

CHAPTER 5

THE CONSONANTAL PHONEMES OF KARANGA

5.0. Introduction

It is the usual practice in the Phonetics literature to describe speech sounds in groups belonging to particular natural classes and, as such, sharing a number of characteristic articulatory, auditory and acoustic features. In this study, as the preceding two chapters will have shown, I shall not depart from this practice. Accordingly, each of the following consonant natural classes will have a separate section: Plosives, Fricatives, Nasals, Affricates, Naso-orals, Rolled /r/ and the labialised phonemes.

The description of each natural class will be prefaced by a brief summary of some of the relevant acoustic work that has been done by other workers in the field. The spectrograms and amplitude sections are then analysed, pointing out the similarities and differences between the consonants of Karanga and those of other languages.

A number of the consonants under study are lingual articulations, that is, they are pronounced by some part of the tongue coming into contact with the hard palate, the alveolar ridge or the dental ridge. For all these consonants some Palatograms were made using the indirect method of palatography. This is a method in which an artificial palate is made for the individual whose speech is to be investigated. Such an artificial

palate is then covered with a marking medium and the palate inserted very carefully into the mouth. The utterance is then pronounced.

The marking medium will be wiped off wherever the tongue has made contact. The artificial palate is removed and a permanent photographic record is made. The medium used in this study is French chalk, and the author was the subject.

Instrumental techniques, such as palatography, are a phonetician's tools, whose function is to supplement his natural tool, the ear. D.M. Beach¹ realised as long ago as 1938 that

"... a trained ear must always be the principal stock-in-trade of the phonetician... The ear method is certainly quicker and generally far more convenient than analysis by means of instruments. It may be called the backbone of phonetic research."

Henry Sweet (1911)² was even more sceptical of instrumental techniques:

"Instrumental phonetics is, strictly speaking, not phonetics at all. It is only a help... The final arbiter in all phonetic questions is the trained ear of the practical phonetician."

However, I must agree with Ladefoged (1964, p.XVI)³

"But for those of us who are not as skilled as Sweet, instrumental phonetics may be a powerful aid and a great use in providing objective records on the basis of which we

1. Beach, D.M. (1938) op.cit., p.11.

2. Sweet, Henry (1911) "Phonetics" in Encyclopaedia Britannica, 11th Edn., Vol.21, pp.1910-1911.

3. Ladefoged, P. (1964) A Phonetic Study of West African Languages, Cambridge.

may verify or amend our subjective impression."

And I shall interpret the palatograms in this study along the lines suggested by Firth (1957)⁴.

5.1. Choosing Examples for Palatography.

In order to establish the points of contact in the pronunciation of lingual consonants, word palatography was used. Palatograms were made to show the effect of pronouncing a whole word as a complete utterance. And, as Firth (1957, p.155) says, "These word-palatograms do not emphasise the sequence of sounds in the utterance of a word. They present features abstracted from the whole utterance."

For this exercise, I chose words containing the sounds in question in syllable initial position before high, front or back vowels. Such palatograms have some interference from the high tongue raising for the vowel and were made mainly to compare them with palatograms of the same sounds in words in which they preceded the low, central vowel [a]. Care was taken to see that each articulation would have only one wipe-off, for the articulation under investigation. So for each such sound there are two Palatograms, e.g. (i) a and b, a being the palatogram of the sound before a high vowel. All the examples chosen for the purpose were numbered consecutively from (i) through (xxv):

4. Firth, J.R. (1957) Papers in Linguistics, 1934-1951, O.U.P.

Number	Word(a)	English	Word(b)	English
(i)	sím bá	"strength"	-pás á	"pass"(v.t.)
(ii)	zím bá	"mansion"	-páz á	"destroy"
(iii)	sím bó	"knobkerry"	-sá wá	"dry" or "wilt"
(iv)	zíp ó	"gifts"	zám á b á	"what you have stolen"
(v)	fúm bá	"lion"	-fám bá	"wash"
(vi)	zè wé	"ear(s)"	-zám bá	"cry hysterically"
(vii)	-d í w á	"be feared"	a í a	"(s)he has panicked"
(viii)	-b í w á	"be eaten"	a b á	"(s)he has eaten"
(ix)	-téma	"cut" or "hit with object"	-tám á	"move house"
(x)	d í b á	"dip-tank"	d á d á	"duck"(n.5)
(xi)	d í m á	"darkness" of night	dám bá	"fruit"
(xii)	ndipe	"give me.."	mánd á	"animal fat"
(xiii)	-pón á	"survive" or "give birth"	-pán á	"give to each other"
(xiv)	jím ó	"bambara groundnut"	-jám á	"meat"
(xv)	nzím b é	sugar cane	-wán z á	"increase in quantity"
(xvi)	nzím b ó	"place"(n.9)	n z á	ideophone of carrying on shoulder.
(xvii)	n b ó	"greediness"	n b á ng á	"grain store"
(xviii)	tsím é	"well"(n.5)	tsám bá	"letter"
(xix)	d z í m bá	"houses"	-d z á m á	"be profound"
(xx)	ít s á	"burn!"	mát s á	"new ones"cl.6.
(xxi)	íd z á	"stumble!"	mád z á	"huge thighs"
(xxii)	t s í p ó	"gift"	t s á p á	"careless person"
(xxiii)	mù d z í ng á of a hill or mountain	"at the foot of a hill or mountain"	d z á w á	"light brown" adj cl.5.
(xxiv)	nd z í w á	"dove"	mánd z á	"applause!"
(xxv)	-rím á	"farm" or "Plough"(v.t./i.)	mà r à r à	"dirt" or "refuse".

5.2. The Palatogram Figure

Firth (1957, pp.148-155)⁵ divided the Palatogram figure into a number of zones in order to make reference to various parts of the mouth where articulatory contact is made easy. His Palatogram figure is reproduced below as Fig. 11:

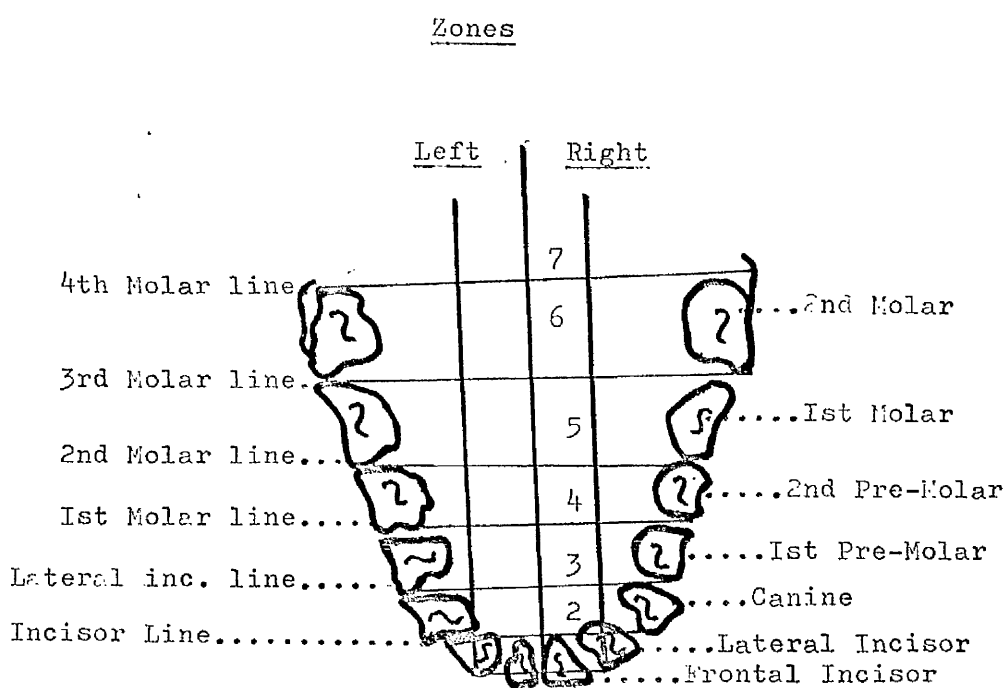


Fig. 11. The Palatogram Figure.⁶

Table III shows the zones of the Palatogram Figure:

5. Firth, J.R. (1957) op.cit.

6. The Palatogram Figure was reproduced on transparent paper by Mr. Stone of the Phonetics Laboratory, SOAS. To find out which part of the mouth a sound is produced in, the Palatogram grid is superposed on the appropriate Palatogram.

The horizontal lines	The Zones	The grouped zones
1. Incisor line	Dental	} Dental
2. Lateral incisor line	Denti-alveolar	
3. Canine line	Alveolar	} Alveolar
4. First molar line	Post-alveolar	
5. Second molar line	Pre-palatal	} Palatal
6. Third molar line	Mid-palatal	
7. Fourth molar line	Post-palatal	
The vertical lines	The vertical zones	
The Median Line	Right and left zones: to speaker's right and left of the median line.	
The Right Line	Right alveolar zone: to the right of the right line.	
The Left Line	Left alveolar zone: to the left of the left line.	

Table III Zones of the Palatogram Figure.

This table explains the way the Palatogram figure above has been "zoned". The zoning is based on the dentition plan. Firth (1957, pp.151-152) added "a natural median line and left and right lines parallel to the median line starting from the interstices between the front and lateral incisors."

5.3. The Plosive Phonemes, Introduction

To recapitulate what has been said in an earlier chapter (chapter 1, pp. 77-79), Karanga has a set of eight plosive phonemes, six of them produced with a pulmonic egressive, and two with a glottalic ingressive, air-stream mechanism:-

pulmonic egressive: /p,ḅ; t,ḁ; k,ḡ/.

glottalic ingressive: /ḅ,ḁ/.

These plosives can be subdivided into three categories depending on their characteristic vocal fold configuration:

voiceless: /p,t,k/.

voiced: /ḅ,ḁ/.

breathy-voiced: /ḅ,ḁ,ḡ/.

Also, depending on their place of articulation, these plosives can further be divided into another three categories:

Front: /p,ḅ,ḅ/, that is, the bilabials;

Central: /t,ḁ,ḁ/, that is, the alveolars;

Back, /k,ḡ/, that is, the velars.

These three interlocking groupings are based on physiological, that is, articulatory, phonatory and auditory criteria alone. An important part of what follows will be devoted to isolating those acoustic features which correspond to these physiological phenomena.

As is apparent from the notation used, the plosives under discussion are distinct phonemes of Karanga: they are found in the words of the language which are semantically differentiated by having one of the plosives at a particular place instead of another. In Central Shona, this place is always the syllable initial. The following are some of the minimal pairs (as such contrastive words are called) in which these plosives occur:

/-ḅárá/ "write" vs. /-pára-, "scrape":- /ḅ/, /p/.

/-ḅárá/, "write" vs. /-ḅárá/, "give birth":- /ḅ/.

/ḍàḍà/, "duck" vs. /-ḍàḍà/, "be snobbish": -/ḍ/, /ḍ/.
 /-tára/, "draw line" vs. /-kára/, "be greedy": -/t/, /k/.
 /gúru/, "big" adj.cl.5 vs. /kúru/, "big" adj.cl.6: -/g/.

On a broader plane plosives belong to the general category of sounds which we have called NON-CONTINUANTS in Chapter 2, that is, sounds during the pronunciation of which air has no exit from the speech tract for an appreciable space of time. This interval of time corresponds to silence on the auditory level and a blank on the acoustic spectrum for voiceless sounds, and to voice, both auditory and acoustic, for the plain-voiced and breathy-voiced sounds (see Spectrograms 47-62).

Research for this class of sounds was done as follows: each of the plosives under investigation was placed in an utterance in which it preceded each of the five vowel phonemes of Karanga. The utterances were then recorded and spectrograms obtained. In order to find out the acoustic correlates of the initiatory dichotomy implosive, [ḍ, ḍ], versus explosive, [p, b; t, ḍ, k, g], a number of these utterances was also analysed using the Electro-Aerometer. The focus of attention here was the behaviour of the stylus both during the hold stage of the plosive articulation, and on the release of the closure for implosives and explosives.

Only three of these plosives have 'lingual' or coronal articulation, namely, /t, ḍ, ḍ/. The description of the articulation of these sounds is supplemented with some palatograms.

5.3.1 The Search for Acoustic Cues.

To put this study into the perspective of phonetic studies generally, I have found it necessary to preface the detailed description of the consonants in each natural class with a summary of the pertinent acoustic research which has been done by other scholars. Most such research has been concerned primarily with the sounds of Indo-European languages. The important differences between the acoustic features of these languages and those of Karanga will be noted, as will the similarities.

Halle, Hughes and Radley (1957)⁷ studied the acoustic characteristics of plosive consonants in some detail. They were concerned to find the most essential cues for the perception of plosives using synthetic speech. They prepared detailed energy density spectra of the noise bursts of plosives. And they also studied the formant transitions of the vowels adjacent to the plosives in a number of words. The efficacy of these two types of cue was assessed by perceptual tests in which isolated segments that contained either stop

7. Halle, M., Hughes, G.W., and Radley, J- P.A. (1957) "Acoustic Properties of Stop Consonants" Journal of The Acoustical Society of America, 29,1,pp.107-116.

burst⁸, or vowel formant transitions alone were present.

As for the cues for distinguishing [p,t,k] from [b,d,g], Halle, Hughes and Radley found that although in English these classes are sometimes distinguished by voice differences, voicing is not a primary distinction. The two sets are distinguished more by aspiration, present in the voiceless set and not in the voiced; and by differences in the duration of the vowel preceding as well as in the spectra of the bursts of the two classes than by voicing.

In addition to the noise burst and vowel formant transition cues, silence was found to be important.

"The acoustic correlates of the complex movements involved in the production of stops are rapid changes in the short-term energy spectrum preceded or followed by a fairly long period (of the order of several centiseconds) during which there is no energy in the bands above the voicing component (above 300Hz)."

-
8. I am aware of Fant's (1973, pp. 111-113) disagreement with this "terminology of the classical Haskins' synthesis" which treats what he calls "transient", "friction" and "aspiration" as a single segment called the "burst". "A fully developed unvoiced stop in stressed, initial position can be decomposed into five successive segments... These are
- (i) occlusion, voiced or silent.
 - (ii) transient. This is the response of the vocal tract to the pressure release exclusive of any turbulence effects.
 - (iii) Fricative segment. This is characterised by noise produced at the consonantal constriction as in a homorganic fricative.
 - (iv) Aspirative segment. This is characterised by an 'h-like' noise originating from a random source at the glottis or from a supra-glottal source at a relatively wide constriction exciting all formants.
 - (v) The initial part of a following voiced sound to the extent that it is influenced by coarticulation with the stop."

Fant, G (1973) Speech Sounds and Features, M.I.T. Press. I shall be content to use the term "burst", subsuming Fant's segments (ii)-(iv) above, noting them individually if they are prominent.

If this acoustic "silence" is filled in with any other type of sound except voice, then a stop will not be perceived. When a stop is adjacent to a vowel all three cues are operative:

Silence
burst
transition } if the stop consonant is initial in a CV-syllable.

and

transition
silence
burst } if the stop consonant is in a post-vocalic position, i.e., -VC.

In Karanga, with an open syllable structure, we shall focus attention on CV-syllables since this will be more informative.

Of these three cues, only "silence", say Halle, Hughes and Radley, is an essential cue - the silence with either transition or noise burst, is a sufficient cue for identifying a stop:

silence + $\left\{ \begin{array}{l} \text{transition} \\ \text{burst} \end{array} \right\} \longrightarrow \text{plosive}$

Thus, for example, in the English word "task", the identification of the final plosive must be attributed to the spectral properties of the burst and "silence": since a fricative intervenes between the vowel and the plosive, the vowel transitions cannot help. The silence tells us there is a stop, and the spectral properties of the burst tell us which stop.

On the other hand, in the ordinary pronunciation of "tact", there is only one "silence" followed by a single burst, although two plosives, [k] and [t] are perceived. The cue for [k] lies in the formant transitions

from the preceding vowel; whereas that for [t] is in the acoustic character of the noise burst.

When they examined the spectra of the noise bursts of the various plosives, Halle, Hughes and Radley found that the three classes of plosives associated with the different places of articulation differed from each other in the following manner:

- (i) [p,b], the labial plosives had a primary concentration of energy in the low frequencies: 500Hz - 1.5KHz;
- (ii) [t,d], the post-dental plosives, had either a flat spectrum or one in which the higher frequencies above 4KHz predominated, aside from an energy concentration in the region of 500Hz;
- (iii) [k,g], the palatal or velar plosives (depending on whether the following vowel is front or back) showed strong concentrations of energy in the intermediate frequency range, 1.5KHz - 4KHz.

According to O'Connor (1973, pp.89-90)⁹, the time factor has to be taken into account in characterising the sounds of speech in general, but it is crucial in the specification of plosives in particular because

- (a) the plosion of a stop corresponds to a burst of noise similar to the noise in fricatives, but short in duration. Because the noise is short, we hear it as plosion; if its duration is extended, we hear it more and more as a fricative. Duration, then, is an essential differentiator between a fricative and a homorganic plosive. Fant (1960,

9. O'Connor, J.D. (1973) op.cit.

p.207)¹⁰ explains the duration of the noise burst thus, "In a prevocalic position, the duration of the noise interval may be defined genetically as the distance in time from the break of contact to the onset of the vocal cord vibration."

(b) the time between the burst and the onset of the following sound is important. After the burst and before the vowel is fully established, a change takes place: the vowel formants bend, that is, they increase or decrease in frequency. This change (formant transition) corresponds to the movement of the speech organs from the stop position to the open position for the vowel.

Ladefoged (1975, pp.177-8)¹¹ makes two interesting points. First, the apparent point of origin of the formant transition for each place of articulation, the locus, is not fixed in frequency; this depends on the adjacent vowel: because the position of the part of the tongue not involved in the pronunciation of the plosive is that for the adjacent vowel (see [ʷ] above).

Secondly, the coming together of F2 and F3 vowel transitions is a distinguishing feature of velar articulations (see Spectrogram 39, [mówónèkà]; and my remarks on pp191-2 above). He also claims that the formant transitions of vowels in the vicinity of velars take longer to accomplish than elsewhere. Presumably this means that the back of the tongue moves less rapidly

10. Fant, G (1960) Acoustic Theory of Speech Production, The Hague: Mouton.

11. Ladefoged, P. (1975) op.cit.

than the lips or the tip and front of the tongue.

The main thrust of these studies seems to be that a satisfactory specification of plosive consonants¹² should include information on:

- (i) the spectral properties of the noise burst: intensity, duration, and its location on the frequency axis;
- (ii) the vowel formant transitions as indicators of place (F2, F3) of articulation;
- (iii) the duration of the acoustic "silence" and whether or not there is voice during the "silence".

In the following description, information will also be sought concerning the articulatory and auditory correlates of the acoustic features observed:

- (i) the place of articulation, with palatograms where appropriate;
- (ii) the air-stream used: whether it is egressive or ingressive, and where initiated;
- (iii) the laryngeal strictures for voicelessness, plain-voice and breathy-voice.

5.3.2. The plosive phonemes.

To obtain acoustic information on the sounds under investigation a number of words were recorded, and analysed using the sound spectrograph. For each of the

12. For a consonant study to be complete, it is sometimes necessary to describe the formant transitions of all the vowels in the language in the environment of each consonant. In this study, however, it has been found that enough information is obtained by investigating the spectra of each consonant in combination with at least one front vowel and one back vowel.

six explosive sounds two spectrograms were obtained, one using a narrow band filter and the other a broad band one. The narrow band filter is helpful in that it is easy to see where the vowel begins by the clear harmonics on the spectrogram. In addition to the two spectrograms an amplitude section of the noise burst was also made on a narrow band filter spectrogram for each plosive. The second advantage provided by the narrow band spectrogram is that it makes it easy to distinguish between those plosives which have voice during the acoustic silence (there are harmonics on the "voice-bar") and those which have voicelessness (there is a complete blank at the base of the spectrum around 250Hz).

On the list of examples below the amplitude section of each sound bears the same number as the narrow band spectrogram since the same utterance was used for both:

<u>Spectrogram before [i]</u>	<u>No. of Section before [i]</u>	<u>Spectrogram of sound before rounded vowel.</u>
47. [bìzǎ́], "horse"	47	48. [bòdòrò], "bottle"
49. [-písá], "burn" (v.t. 1.)	49	50. [-púrá], "thrash" (v.t.)
51. [dìbà], "dip-tank"	51	52. [dòróbà], "town"
53. [-tízǎ́], "escape" (v.i.)	53	54. [-túrá], "lay down burden"
55. [gǐjà], "boast" (v.i.)	55	56. [gúrú], "big" (adj. cl.5)
57. [kítsí], "cat"	57	58. [-kúrá], "grow up".

Since the implosives do not have a noise burst on the spectrum, no amplitude sections are provided for them:

only two spectrograms for each, one before a front,
and the other before a back, vowel.

59. [-ḡisá], "remove"

60. [-ḡurá], "remove pot
from fire"

61. [-ḡísá], "like very much"

62. [-ḡurá], "confess".

In order to observe the difference between implosives
and explosives more, Mingograms were made for the
following utterances:

1. [ḡòróḡà], "town";
2. [ḡáni], "plain, flat country";
3. [ḡòrò], "beer".

5.3.3. The bilabial explosives, /p, b/.

(a) Pronunciation (Diagram 8 p.217).

To produce these sounds, the two lips come
together to form an air-tight closure. The soft palate
is raised, cutting off the nasal cavity. Since the
tongue is not involved in pronouncing these sounds, it
assumes its position for the adjacent, most probably
the following, vowel.

The bilabial closure is maintained for a period
of time. During this passage of time, pulmonic air
passes through the glottis into the supra-glottal
cavities, except the nasal cavity. This air has two
modes of flow in these cavities: aperiodic and periodic.
In the case of [p], the vocal folds are wide apart
throughout their entire length, allowing a volume of
aperiodically vibrating air into the pharynx-mouth chamber.
For the breathy-voiced [b] on the other hand, there is

simultaneously periodically and aperiodically vibrating air. According to Ladefoged (1971, pp.6,8)¹³ to produce breathy-voice or murmur, the arytenoids are held apart, while the ligamental vocal cords are vibrating. Overall, this stricture allows less air to enter the supra-glottal cavities than that for [p].

Since the labial and velic closures are complete, the air-flow into the pharynx and mouth chamber, whose capacity is finite, cannot go on beyond a certain point. The bilabial closure forces pulmonic air to be compressed until its pressure is higher than that of atmospheric air. There is an explosion when the two lips part suddenly and the two air pressures are balanced up. The release of the bilabial closure is always to a wide degree of approximation for plosives. This contrasts with the release for affricates, which is to a narrow degree of openness, producing homorganic friction.

The explosion has a sharp ring for [p], but is essentially dull and throaty for the breathy-voiced [b]. For both, pulmonic air has a median mode of egress from the speech tract.

(b) Acoustic Features

In order to compare the amplitude sections of the noise bursts of any two of these plosives, I made a tracing of the section of one on transparent paper and then superimposed the tracing on to the section of the second burst. For [p] and [b] in particular, the

13. Ladefoged, P. (1971) op.cit.

amplitude section of the voiceless sound has much more energy than that of the breathy-voiced one.

The overall amplitude profiles are, however, similar, with peaks of energy at the same points along the frequency axis - see amplitude sections 47 and 49. The second difference between these amplitude sections is that the voiceless plosive not only has higher amplitude peaks, but it also has an energy concentration in the high frequencies greater than that for the breathy-voiced sound. The noise burst of the voiceless sound starts virtually on the base line and extends as high as 5 KHz (spectrogram 49); but that of the breathy-voiced [b] has a short delay and begins at about 250Hz and is very faint - see spectrograms 47 and 48.

During the hold stage for the plosives, there is a blank on the "voice-bar" (80-250Hz) for [p] both on spectrogram 49 and spectrogram 50. The breathy-voiced [b] has some striations on the voice-bar on the broad band spectrogram 48. These striations are rather far apart and are concentrated only in the first half of the hold stage. Both points are confirmed on spectrogram 47, the narrow band spectrogram. The striations which seem far apart on spectrogram 48 in fact represent only two low-intensity harmonics; and, secondly, there is a complete break between the two harmonics and the first two harmonics of the ensuing vowel. The noise burst which follows the release of the labial closure begins just above the first harmonic of the stop and the vowel.

The second formants of the vowels [o] and [u]

preceded by [b̥] and [p] respectively on spectrograms 48 and 50 have initial transitions pointing towards a very low locus, probably at 600Hz.

There appear to be no dramatic differences generally, between the sounds investigated by Halle, Hughes and Radley (1957) and the plosives [b̥] and [p], apart from the minimal voicing for [b̥] and its consequent guttural auditory impression. O'Connor's point about the difference between plosives and fricatives being partly attributable to the difference in the duration of the fricative portion of the two sound types is confirmed: the voiceless Karanga plosive [p] has a noise burst with a duration of about 5 centi-seconds and that of [b̥] is even shorter. Perhaps the greatest difference between Halle, Hughes and Radley's bilabial sounds and those of Karanga lies in the upper limit, on the frequency axis, of the noise component: it seems to be considerably higher for Karanga plosives, whether a front or a back vowel is following. But, as we shall see in the next section, the amplitude of this noise burst is uniformly low compared to that for the alveolar plosives.

Finally, whereas for English plosives Halle, Hughes and Radley found the dichotomy voiced:voiceless less preferable to lax:tense, in Karanga the dichotomy breathy-voiced:voiceless is operative even in utterance initial position.

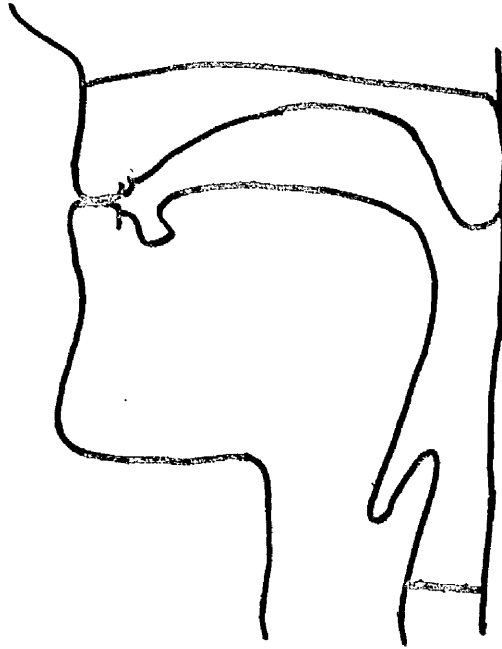


Diagram 8. The bilabial plosives [p, b].

5.3.4. The alveolar explosives, /t, d/.

(a) Pronunciation

There are four Palatograms [(ix) a, b and (x) a, b] to be taken into account, two for each plosive: one before a front, high vowel and the other before a low vowel.

For purposes of comparison, the examples (x) a and b will be used, that is, [d̪íɸà] and [d̪àd̪à]. Both Palatograms show that these two explosives are alveolar articulations. There is firm contact between the sides of the tongue and the left and right alveolar zones. The contact starts at the extreme back of the palate, behind the fourth molar line, and extends as far forward in the mouth as the canine line. The central area contact

is as wide as both the left and right central zones together. It starts from about the first molar line and stops just short of the lateral incisor line.

The difference between these two realisations of [d] lies in the width of the contact between the sides of the tongue and the left and right alveolar zones. This contact has a width which averages 16mm when the sound is followed by the front high vowel [i] in [dɪpə]. When the "same" sound precedes the open, non-back and non-front vowel [a], the width of the contact between the sides of the tongue and left and right alveolar zones averages 11.5mm. The difference between these two measurements is accounted for by the

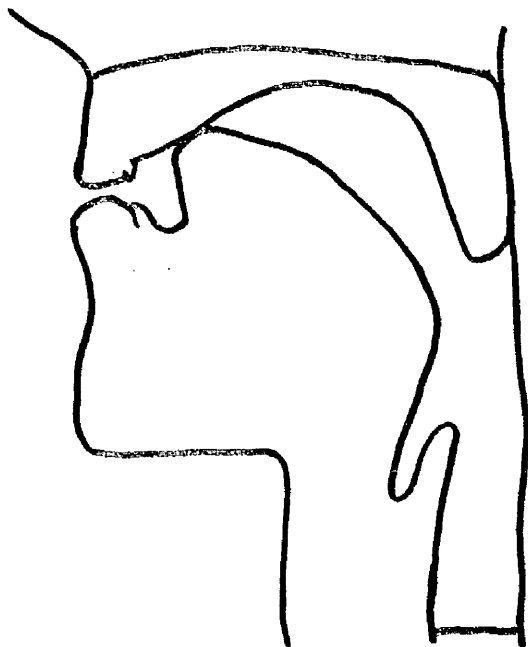


Diagram 9: The alveolar plosives [t, d].

amount of interference from the raised front of the tongue in anticipation of its position for the vowel [i]. There is no such raising of the tongue in the pronunciation of [d̥] in Palatogram (x) b [d̥a̠a̠].

These features are comparable to those observed on Palatograms (ix) a and b, [téma] and [táma] respectively. Such small differences as may be noted will be due to the fact that the tongue is raised higher for [i] than for [e] in [d̥i̠a̠] and [téma].

Throughout the pronunciation of [t], the vocal cords are held wide apart, allowing a strong current of aperiodically vibrating air to flow from the lungs into the supraglottal cavities. But for [d̥], the ligamental vocal cords vibrate periodically, while the arytenoids are kept wide open: as for breathy-voice or murmur (see Ladefoged (1971, pp.6,8)).

The soft palate is kept raised to exclude the nasal chamber. Pulmonic air is thus channelled into the oral cavity, where it is compressed behind the alveolar closure. When this closure is suddenly released, plosion is heard as the compressed air from the speech tract balances up its pressure with that of atmospheric air.

(b) Acoustic Features

A comparison of the amplitude sections of the noise bursts of [d̥], 51, and [t], 53. shows three outstanding similarities: first, that the noise covers virtually the whole frequency range, beginning from the base line to over 8KHz; and, secondly, that the amplitude

peaks are at corresponding points along the frequency axis, which shows that the two sounds have the same supra-glottal cavity configuration; and, lastly, although it starts in the very low frequencies, the noise burst for each sound has its highest amplitude in the high frequencies above 4KHz. This is a feature which, as Halle, Hughes and Radley (1957) say in their paper, distinguishes the alveolar plosives from the other two classes.

Spectrograms 52 and 54 [d̀òró̀b̀à] and [-túŕá] respectively, were made using a broad band filter. The first spectrogram shows that there is some voicing during the hold stage for [d̀], although the striations are wide apart and die away towards the end of the hold stage, the acoustic "silence". The second spectrogram shows that [t] has no voicing during the hold stage, although the base line is not too clearly defined.

The narrow band spectrograms 51 and 53, [d̀í̀b̀à] and [-tí̀z̀á] respectively, on the other hand, show that the far-apart and low-intensity striations in the voice-bar of [d̀] correspond to three harmonics (spectrogram 51), and the ill-defined base line for [t] has a blank above it. On spectrogram 51, as on 53, the noise burst for both sounds seems to start on the base line: for [d̀] there is a very short interval between the end of voicing for the plosive and the noise burst.

The duration of the noise burst is more clearly visible on the narrow band spectrograms than on the broad band ones, mainly because it is easier to see where the random noise ends and the harmonics of the vowel begin:

the burst of [t] has a duration of about 6 centiseconds; while that of [d] seems to be less than half of that, at the base line.

On all the four spectrograms of these sounds the second and third formants of the vowels adjacent to the plosives have transitions pointing towards higher loci than for the bilabials above. The utterance [dòróbà], spectrogram 52, provides useful contrast between [d], the alveolar, and [b], the bilabial breathy-voiced stops: both are adjacent to the rounded vowel [o]. The alveolar sound precedes this vowel whose initial F2 transition bends upwards, pointing towards a high locus at about 1.7KHz. The third formant of the same vowel points towards a locus which is slightly higher than that, at about 2KHz. By contrast the "same" vowel preceding [b] in the same utterance has terminal F2 and F3 transitions pointing towards much lower loci: in fact both these formants bend downwards towards the loci for [b] at about 500Hz and 700Hz respectively. On spectrogram 54 the loci of F2 and F3 of the transitions of vowel [u] preceded by [t] seem to be as high as those of [o] in the environment of [d] above.

The auditory correlates of these acoustic cues for [d,t] are the same as for [b,p] above: the breathy-voiced sound has a dull, throaty plosion, whereas the unvoiced counterpart has a clear-cut, sharp ring.

5.3.5. The velar explosives /k,g/.

(a) Pronunciation.

To produce these sounds, the back of the tongue is raised until it makes firm contact with the soft palate. At the same time there is a velic closure, cutting off nasal resonance. The position of the lips depends on that of the adjacent, particularly the following, vowel. There is, by kinaesthetic judgement, a difference in the actual point of contact between the back of the tongue and the soft palate, depending on whether the following vowel is front or back or mid. The point of contact is farther in the mouth, near to the hard palate, when the following vowel is front and high. As we shall see below, this allophony has a bearing on the acoustic features observed for [k] and [g].

During the pronunciation of [k], the voiceless sound, the vocal folds are wide apart, allowing aperiodically vibrating air to flow from the lungs into the pharynx-mouth chamber. Less, but largely periodically vibrating air is allowed through during the pronunciation of the breathy-voiced [g]. Ladefoged (1971 ,pp.13-14)¹⁴ says of the murmured or breathy-voiced sounds of African languages (including Shona) "... during the murmured sounds the vocal cords seem to me to be held slightly closer together than in the Indian languages, so that there is more voice and less breath escaping."

14. Ladefoged, P. (1971) op.cit.

In both cases the break of the velar closure to release the compressed air is radical, that is, to a wide degree of openness, producing no audible homorganic friction. In Karanga (and Central Shona generally) there are no affricates involving these velar explosives. This means that the release of the velar closure is always followed by plosion, sharp and clear-cut for the unvoiced [k] and dull and throaty for the breathy-voiced [g].

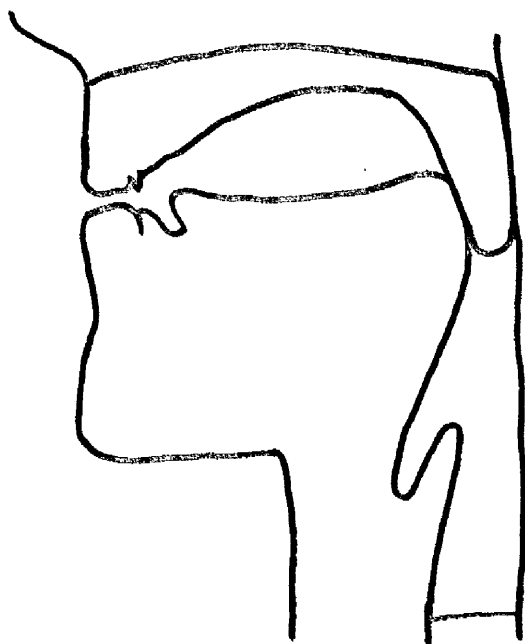


Diagram 10 The velar explosives /k,g/.

(b) Acoustic Features

The amplitude sections of the noise bursts of these sounds (see amplitude sections 55 and 57) have the following similarities: first, there is some energy spread from the lowest frequencies to the highest

frequencies; secondly, this energy has its highest peaks in the middle of the frequency range, between 2.4KHz and 6KHz: the first two points mean that the two amplitude sections have a similar overall profile.

The narrow band spectrograms 55 and 57 [gíjǎ] and [kítsí] respectively show that this noise burst has longer duration on the latter spectrogram, that is, for [k], than on the former: about 6 centi-seconds and 3 centi-seconds. But both are, as expected, shorter in duration than the fricative components of the hom-organic sounds [x, ʃ].

On the broad band spectrograms 56 [gúrú] and 58 [-kúrǎ], the vowels immediately following [g] and [k] have no energy above 1KHz. But the first spectrogram shows that there is voicing during the hold stage of [g], indicated by the wide-apart striations. The unvoiced sound [k] has no such striations. On spectrogram 55 the voicing for [g] is represented by three low-intensity harmonics.

Since the broad band spectrograms 56 and 58 do not provide any information concerning formant transitions of vowels adjacent to [g] and [k], we turn to spectrograms 89 and 90, [-nàkǎ], "be nice or sweet" and [nǎkǎ], "inheritance", respectively. On these spectrograms, the voiceless velar plosive [k] is flanked by two realizations of the "same" vowel [a]. It is remarkable how the F2 and F3 transitions of the vowel bend towards a common locus, both before and after the sound [k] in both utterances. This locus is at about 1.9KHz, which is lower than that which the transitions have when [k], or,

for that matter [g], is adjacent to a front high vowel: see the narrow band spectrograms 55 and 57, where the transitions of the formants of [i] seem to have a locus of at least 2KHz. This finding corroborates what Delattre (1968)¹⁵, among others, reports: the point of contact for velar articulations shifts depending on whether the adjacent vowel is front or back. This means that here we have very close to palatal [k] and [g]. And, what is more important, this accounts for the difference in the frequency range over which the noise burst of those sounds is concentrated on spectrograms 55 and 57 (i.e. 2.4KHz to 6KHz) on the one hand, and on spectrograms 56 and 58 (i.e. 500Hz to 3KHz, where it is very weak indeed) on the other.

15. Delattre, P. (1968) op. cit.

5.3.6. The bilabial implosive, /ɓ/.

(a) Pronunciation.

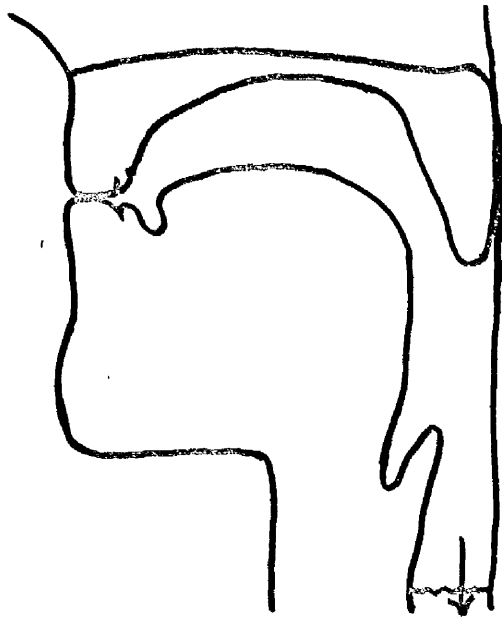


Diagram 11, /ɓ/.

As the diagram above shows, to produce this sound, the two lips come together, making firm contact as for the explosive [p]. Diagram 11 also shows that the soft palate is raised to make a velic closure, cutting off nasal resonance. The position of the tongue is that for the following vowel.

An ingressive air-stream mechanism is used in the pronunciation of this sound. The vocal cords are positioned for producing normal voice, and the larynx as a whole moves downwards - residual lung air making the vocal cords vibrate periodically. This downward movement of the larynx increases the capacity of the supra-glottal cavity, the pharynx-mouth cavity. This creates "negative" air pressure, to use Catford's (1970)¹⁶

16. Catford, J.C. (1970) op.cit.

term, that is, the air in the pharynx-mouth cavity is rarefied. Its pressure being lower than that of atmospheric air, a potential or actual inward flow of air into the speech tract is thus created.

The lip closure is then released, allowing atmospheric air to rush into the speech tract. This rush of air produces a hollow pop as the air pressures are balanced up.

(b) Acoustic Features

Spectrograms 59 [6¹isá] and 60 [6¹urá] show that during the closure or acoustic "silence" stage for [6], there is vocal cord vibration - a low frequency voice-bar with its centre at about 250Hz, and with a duration of 15 centi-seconds on spectrogram 59. The striations which correspond to this voicing are weak at the beginning, before the vocal cords have attained a steady-state of vibration: but the latter part of the voice bar has striations with bolder outlines.

Unlike the explosives discussed above, [6] does not seem to have a noise burst - hardly likely since air is sucked into the vocal tract on the release of the bilabial closure. The acoustic correlate of the auditory "hollow pop" can be seen on a M⁶ongogram tracing (below).

The vowel [i] on spectrogram 59 has initial F2 and F3 frequency transitions pointing towards low loci between 500Hz and 1KHz. These transitions are very short and rapid, compared to the terminal transitions

of the same vowel before the dental fricative [s].
On spectrogram 60 [bùrá], where [b] precedes the back rounded vowel [u], with a low second formant, there are no clear formant transitions.

So far, two acoustic (as opposed to auditory) differences have been noted between the explosive [b] and the implosive [ɓ]: the presence of a noise burst on the spectrum of the former, missing on that of the latter; and the continuation of the voice-bar of [ɓ] into that of the following vowel, whereas there is a break between that of [b] and the ensuing vowel.

The Mingograms of recorded utterances 59 [bisa'] and 47 [biza'] did not display any spectacular differences between [ɓ] and [b] apart from the oscillations of the former sound being closer together than those of the latter.

So some utterances spoken directly into the Electro-Aerometer, thereby providing information on two additional dimensions - the oral and nasal tracings - were analysed in order to see if any more acoustic differences exist between the explosive sound and its implosive counterpart. Mingograms 1 [dòróhà], "town" and 3. [bani], "plain" where [b] and [ɓ] both precede the "same" vowel, were obtained. Mingogram 3 shows that during the closure stage the mouth stylus draws a falling tracing for [ɓ], corresponding to the falling pressure of the air in the pharynx-mouth chamber. During the corresponding stage for [b] on Mingogram 1, the stylus draws a fairly level tracing. When the closure is released the stylus catapults upwards for the explosive, and down-

wards for the implosive. Some voicing continues during the downward movement of the stylus for [ɓ], but for [ɓ̥], the upward movement of the stylus is represented by a straight line: which corresponds to the break in the voice-bar between [ɓ̥] and the following vowel on the spectrograms, not observed for [ɓ]. On these Mingograms both the linear and the logarithmic intensity scales show that whereas intensity rises suddenly following the release of the bilabial closure for [ɓ], it falls for [ɓ̥].

5.3.7. The alveolar implosive /ɗ/.

(a) Pronunciation.

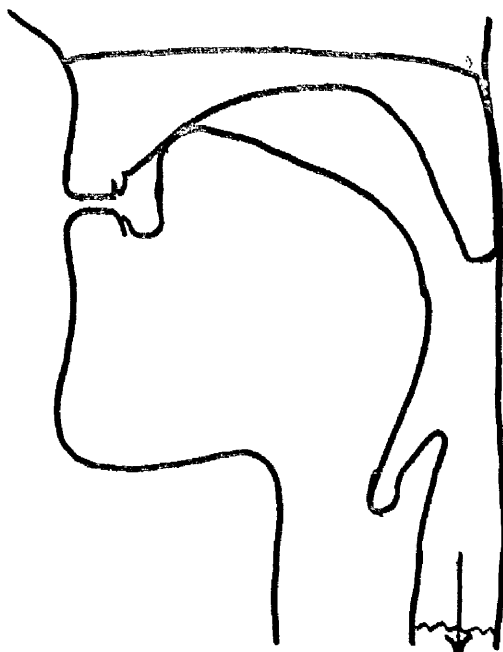


Diagram 12. The alveolar implosive /ɗ/.

The alveolar closure for this sound is the same as that for its explosive counterpart. Two Palatograms (xi) a and b were made for the alveolar implosive, one in the context of the high front vowel [i] and the other in that of the open vowel [a] in the words [dímà] and [dàmbá].

As Palatogram (xi) a shows, there is firm contact between the sides of the blade and front of the tongue and the left and right alveolar zones. The strange thing about this particular Palatogram is that although the alveolar implosive is here next to the front high vowel [i], the closure on the sides is very narrow, much narrower than that made when the sound is followed by the open vowel [a] in Palatogram (xi) b.

The contact starts at the very back of the palate, behind the fourth molar line, extending as far forward in the mouth as the middle of the first pre-molar. The difference between the two examples can be seen more clearly when we examine the central area contact, the example a having wider contact than b, especially in areas 4 and 5 of the left central zone.

This central area contact extends as far forward in the mouth as the middle of the area 3, covering both the left and the right central zones, which is within the area covered by alveolar articulations.

The position of the lips is the same as that of the adjacent, especially the following, vowel. The soft palate is raised throughout this articulation, excluding nasal resonance. The same air-stream mechanism as that

used for producing the bilabial implosive is used here. After the body of air enclosed between the alveolar contact and the larynx has been rarefied by lowering the larynx, the alveolar contact is released. This also results in a hollow plosive sound as the enclosed air balances up its pressure with that of atmospheric air.

The vocal cords vibrate periodically as the larynx is lowered through residual lung air.

(b) Acoustic Features.

Spectrograms 61 [dísá] and 62 [dùrá] show that during the hold stage for [d] there is vocal cord vibration on the voice-bar at about 250Hz. This voicing has low-intensity and far-apart striations at the beginning. In the latter part of the voice-bar the striations are closer together and have a bolder outline, continuing into the voice-bar (the F1) of the following vowel without a break. The release of the oral closure is not followed by a noise component as is that of the explosive alveolar above.

The vowel [i] has F2 and F3 initial frequency transitions pointing towards loci at about 1.7KHz and 2KHz respectively. The vowel [u] on Spectrogram 62 also has an initial F2 transition pointing towards the same high frequency area.

Compared to spectrogram 51 [dípà], the striations of [d] seem to be closer together than those of [ɖ], the breathy-voiced explosive. The contrasts between the explosive and implosive bilabials observed above also show

between the alveolars. Mingogram 1 [d̥òróbà] has the explosive alveolar in an utterance initial position. This sound's acoustic features are in contrast with those of [d] on Mingogram 2 [d̥òrò], "beer". During the hold stage for [d̥] both the linear and logarithmic intensity scales show a rise, as during that for [d]. The mouth stylus, by contrast, shows a horizontal tracing for the hold stage of [d̥], but a falling one for [d]. When the mouth closure for these stops is released, the stylus shoots upwards in a straight line for [d̥], but plunges downwards for [d], showing some oscillations.

And although both sounds are initial in a low tone syllable in their respective utterances, the oscillogram shows that the oscillations for the voicing of [d̥] are farther apart and lower in amplitude than those of [d] - reflecting a lower and a higher frequency of vocal cord vibration respectively. Consequently, the pitch of the two syllables has different profiles: that of the [d̥ò-] is low and level; that of [d̥ò-] is high at the beginning and falls progressively until it levels with that of the breathy-voiced syllable.

5.3.8. The alveolar click /ɟ/.

In addition to these two implosive stops, a click made at the alveolar ridge is widely used by Karanga speakers in the Selukwe, Gwelo and Que Que areas of the country. This sound is symbolised by the International Phonetic Association by the symbol [ɟ] and occurs in such words as:

/mùʔìrò/, "whip made of animal skin"	<	Ndebele
/kùʔùbà/, "to rinse mouth"	<	Ndebele
/màʔimbi/, "edible caterpillars"	<	Ndebele

These items are borrowed from Ndebele and, ultimately, from Zulu. The sound in question has not been replaced by one of the mainstream Karanga sounds, unlike the alveolar lateral frictionless continuant of the first example, /l/:

/mùʔìlò/, "whip....", which is regularly replaced by the alveolar rolled consonant /r/.

Pronunciation.

As shown in the diagram 2b Chapter 2 (p.96), firm contact is made between the tip-and-blade of the tongue and the alveolar ridge. There is also a second closure between the back of the tongue and the velum, in addition to the velic closure.

There is a downward and backward movement of the central part of the tongue. This creates negative air pressure (rarefaction) behind the alveolar (articulatory) closure. This creates an actual or potential flow of air into the oral cavity. The release of the alveolar contact is radical and results in a sharp, sucking sound. There is, of course, no voice during the production of [ʔ].

The number of words in which this sound occurs is very small indeed. And most people would not want to consider it as anything more than an adventitious sound, usually used by people who have a certain degree of competence in the source language, Ndebele.

/mùʔìrò/, "whip made of animal skin"	<	Ndebele
/kùʔùbà/, "to rinse mouth"	<	Ndebele
/màʔim̩bĩ/, "edible caterpillars"	<	Ndebele

These items are borrowed from Ndebele and, ultimately, from Zulu. The sound in question has not been replaced by one of the mainstream Karanga sounds, unlike the alveolar lateral frictionless continuant of the first example, /l/:

/mùʔìlò/, "whip....", which is regularly replaced by the alveolar rolled consonant /r/.

Pronunciation.

As shown in the diagram 2b Chapter 2 (p.96), firm contact is made between the tip-and-blade of the tongue and the alveolar ridge. There is also a second closure between the back of the tongue and the velum, in addition to the velic closure.

There is a downward and backward movement of the central part of the tongue. This creates negative air pressure (rarefaction) behind the alveolar (articulatory) closure. This creates an actual or potential flow of air into the oral cavity. The release of the alveolar contact is radical and results in a sharp, sucking sound. There is, of course, no voice during the production of [ʔ].

The number of words in which this sound occurs is very small indeed. And most people would not want to consider it as anything more than an adventitious sound, usually used by people who have a certain degree of competence in the source language, Ndebele.

5.3.8. Conclusion

On the basis of the observations made above, the plosives of Karanga can be grouped into the following classes, previously determined on articulatory criteria alone (p.204):

Articulatory Features

Acoustic Features

Front:[p,ḡ,b]

- (i) Both F2 and F3 transitions of adjacent vowels point towards low loci, between 250Hz and 700Hz.
- (ii) The diffuse spectrum of the noise burst of [p,ḡ] originates from the very low frequencies to about 5KHz.

Central:[t,ḡ,d]

- (i) Both F2 and F3 transitions of adjacent vowels have high loci, between 1.8KHz and 2KHz.
- (ii) The energy of the noise burst is concentrated in the frequency range 4KHz upwards. These two features are shared by the front allophones of [k,ḡ].

Back:[k,ḡ]

- (i) When followed by [a], the F2 and F3 transitions point towards a common locus at about 1.6KHz.
- (ii) The energy of the noise burst is concentrated in the frequency range 500Hz to 3KHz, that is, half-way between [t,ḡ] and [p,ḡ].

The distinction between voiceless, voiced and breathy-voiced plosives in Karanga has a number of acoustic correlates:

Phonatory Features

Acoustic Features

Voiceless: [p,t,k]

- (i) The noise burst has about twice the duration of that of the corresponding breathy-voiced sound.
- (ii) There are no striations on the "voice-bar".
- (iii) The amplitude of the noise burst has higher peaks than that of the breathy-voiced counterpart.

Phonatory FeaturesAcoustic Features

- Breathy-voiced: [b̥, d̥, g̥] (i) The noise component has about half the duration of that of the voiceless counterpart.
- (ii) The "voice-bar" has some striations which are rather far apart. On narrow band spectrograms this corresponds to about 2 harmonics.
- Voiced: [b, d]: (i) There is no noise burst after the release of the oral closure.
- (ii) The striations on the "voice-bar" are closer together than those of the breathy-voiced explosives. These striations continue into the F1 of the following vowel without a break.
- (iii) The dichotomy between implosive and explosive is shown on Mingo-grams by the stylus of the mouth tracing catapulting upwards immediately the oral closure for an explosive is released, and plunging downwards for the implosive.

5.3.9. The Phonemic Interpretation of the Stop Consonants.

In A Comparative Study in Shona Phonetics, as I have already noted, Doke (1931)¹⁷ did not work within the framework of an explicit phonemic theory. I refer to his important publication here because, as we have seen (Doke, 1931, p.92) he regarded breathy-voiced sounds as a complex cluster of Consonant + h. This is one of the phonemic interpretations advanced by Stevick (1960)¹⁸ to whom I shall pay more attention below. Fortune (1955)¹⁹

17. Doke, C.M. (1931) op.cit.

18. Stevick, E.W. (1960) "The Implosive-Explosive Contrast in Manyika" African Studies, 19,2, pp.88-95.

19. Fortune, G (1955) op.cit.

leaning "heavily on Doke's previous work", to use H.A. Gleason's expression, also regarded breathy-voiced sounds as being sequences of C + h.

Although in Shona Grammatical Constructions Fortune (1968) regards such sounds as one complex segment, each with breathy-voice as a component feature, he nevertheless offers no phonetic description of the phenomenon. He, however, presents a phonemic analysis (pp.77-78 above). This in itself represents a radical departure from the common practice in Shona phonetic studies. Stevick's paper is a study of the explosive-implosive dichotomy in the Manyika dialect.

Fortune's phonemic analysis of the plosives of Shona, whereby implosives and explosives are treated as allophones of one phoneme, is only one of a number of possible analyses put forward by Stevick (1960). Stevick is pragmatic, where his predecessors in the study of Shona seem to be content to put forward one view, thereby implying that theirs is the only correct analysis.

On the phonetic level, Stevick recognises three "Ranges":

- (i) the voiced explosive sound with h -quality, i.e. the breathy-voiced sound;
- (ii) the plain voiced implosive;
- (iii) the plain-voiced explosive. Like this author, Stevick does not recognise velar implosives, which Fortune (1968,p.13) says exist in Shona.

The first phonemic interpretation considered by Stevick is one whereby Ranges (i) and (iii) above are grouped together as allophones of one explosive phoneme,

while Range (ii) is treated as one implosive phoneme:

- (a) /ɓ/, with allophones [ɓ̥] in ɓV and NɓV
and [ɓ] in NCV.

(where N stands for a homorganic nasal with the same voice quality as C, the plosive phoneme).

- (b) /ɓ̥/, with allophones in CV,
(where V stands for any of the five vowel phonemes of Central Shona).

The same analysis is extended to the alveolar series, that is:

- (a) /ɗ/, with allophones [ɗ̥] in ɗV and NɗV
and [ɗ] in NCV
(b) /ɗ̥/, with allophones in CV.

This solution, as Stevick notes, is the usual solution in phonemic analysis. The solution entails giving the implosive-explosive dichotomy priority over the presence-absence of breathy-voice.

The second solution offered by Stevick is the same as the one presented by Fortune (1968, pp. 22-24) i.e. pairing up (ii) and (iii) as members of one plain-voiced phoneme, and leaving (i) as one h-quality phoneme.

"This second solution, assigning an explosive and implosive to a single phoneme, neither requires nor permits us to ignore the unquestionable difference which exists between these two types of sound,"

says Stevick. I have called Stevick's approach 'pragmatic' because later on in his article he says not one of the interpretations he offers "may be either defended or attacked on the grounds of adequacy, since all three subsume exactly the same phonetic data." But his comment on the second

phonemic analysis above is not just pragmatic: it is non-committal. It leaves unanswered the question as to the status to be given to "the unquestionable difference between these two types of sound."

The second solution, as Stevick himself admits, demotes the implosive-explosive dichotomy "from the status of distinctive feature" and re-enters it "into our description at a lower level - the level on which we describe the environmentally conditioned allophones of the phonological units." In short, the choice between the first solution and the second depends on what status we decide to give to the two dichotomies: plain-voice vs. breathy-voice on the one hand, and implosive vs. explosive on the other.

Stevick's reference to environmental conditioning here reminds one of the situation in Shona phonetics whereby a plain-voiced bilabial or alveolar plosive in a C'V-syllable has to be implosive: that is, in C'V, if plain-voiced, then implosive. But also, if C is breathy-voiced, then explosive. The latter agrees in voicing with the other consonantal phonemes with which it clusters. And since the language does not have clusters between a voiced or breathy-voiced sound with a voiceless one, the clusters in which a plosive occurs are either wholly plain-voiced or wholly breathy-voiced: [mb or m̥p, by, d̥z, d̥ʒ, nd, n̥d], for example.

Therefore, the two dichotomies above seem to be ultimately environmentally conditioned, or conditioning. The focus of the rest of this discussion will have to be on finding out whether there are any plausible theoretical

grounds for regarding either dichotomy as comprising features which can also be treated as more intrinsic.

Left to myself, I would want to consider the implosiveness of [ɓ,ɗ] as an intrinsic, distinctive feature of the sounds, and not one that is conditioned. Having been introduced to the discipline by way of the study of English structure, I can say that the situation in that language, whereby a voiced sound is partially devoiced in certain environments, e.g. /b/ in #CV- or -VC#, as opposed to VCV, can hardly be equated with what to me seems to be a more radical and fundamental change in Karanga: the choice in English between [b] in VCV and [ɸ] in #CV- or -VC# is allophonic, whereas that in Karanga between [ɓ] an implosive and [ɸ], an explosive, is phonemic. In the latter case two opposite air-stream mechanisms are involved, whereas in the former, there is a slight adjustment in the position of the vocal cords.

Another point worth noting is that among the languages of the world where more than one air-stream is used, and especially those of the African sub-continent (Ndebele, Zulu, Xhosa, etc) Shona would be the only one which treated one of those air-streams as being allophonic, that is, environmentally conditioned, rather than intrinsic and distinctive.

I am therefore persuaded to go along with what Stevick has called "the usual solution", that is, the pairing up of the explosive ranges as allophones of one phoneme, and the treating of the implosive range as a separate phoneme. This interpretation has a number of

things to recommend it. First, I maintain that by giving the implosive vs. explosive dichotomy priority over the plain-voiced vs. breathy-voiced one, this analysis satisfies the criterion of phonetic plausibility more than its counterpart. A change in the source and direction of an air-stream is more radical than one involving the configuration of the vocal cords and should be given distinctive feature status. This also satisfies my own intuitions. To regard the implosive vs. explosive dichotomy as being environmentally conditioned would be counter-intuitive.

5.3.10.0. The Fricative Phonemes.

5.3.10.1. Introduction.

The number of fricative phonemes one uses in Shona depends on one's linguistic background. In my speech, the Mhari sub-dialect of Karanga, there are thirteen fricatives, namely:

/f, ɣ; s, z, ʒ, ʒ̥; ɬ, ɮ; x, ɣ̥; h/.

All these fricatives are produced with an egressive air-stream and all, except /ɬ, ɮ/, have a median oral mode of egress. But like the plosive phonemes above, the fricatives of Karanga can be divided into a number of sub-categories depending on articulatory and phonatory criteria. Taking the latter criterion first, there are two possibilities:

Voiceless: /f, s, ʃ, ɬ, x/.

breathy-voiced: /ɣ, z, ʒ, ʒ̥, ɮ, ɣ̥, h/.

Depending on their place of articulation, these fricatives can be divided into three further categories:

Front: /f, v; s, z/

Central: /ʃ, ʒ; ʧ, ʤ; ʒ̥, ʒ̥̥/

Back: /x, ɣ, ɦ/.

The fricatives /s, z; ʃ, ʒ; ʧ, ʤ; ʒ̥, ʒ̥̥/ also form a class of their own in articulatory terms since they are the so-called "lingual" articulations. The articulatory strictures for all these phonemes are made between parts of the roof of the mouth and those of the tongue, ranging from the tip of the tongue to the front of the tongue. Palatograms (i) through (viii) (see list on pp. 201 above) were made for these sounds.

In order to obtain acoustic information on these fricatives, thirteen utterances in which the fricative sounds were flanked by the same vowels, namely [a-u] were recorded (/ʧ, ʤ/ are exceptions: [u-a], and so are /x, ɣ/: [a-o]).

Broad band spectrograms were obtained and amplitude sections made of the fricative component of each sound under investigation. Below is a numbered list of the examples used. The amplitude section bears the same number as the utterance in which the sound in question occurs:

<u>Numbered Spectrogram</u>	<u>English</u>	<u>Number of Section, and sound.</u>
63. [isá pámafud̥zi]	place (it) on the shoulders.	63. [f]
64. [pānēmáyud̥zi]	there's hair on...	64. [y]
65. [saramúmasurè]	stay behind	65. [s]
66. [pāvēnāmāzúwá]	it's a few days since.	66. [z]
67. [tāvāpākuká]	we are now eating.	67. [k]
68. [unópāpākuká]	you(sing.)fear too much	68. [f]
69. [tām ^w amasūšú]	we have drunk "soft drinks"	69. [ʒ]
70. [ázúwà mèsó]	(s)he blushed	70. [ʒ]
71. [zāarimāfūrā]	this is a bad omen	71. [ʃ]
72. [ndāwānāmāzūm ^w]	I have found 'fruit'	72. [ʒ]
73. [tāxórānāsí]	we've been paid today	73. [x]
74. [ndīnódá ʒowà]	I like mushroom	74. [ʒ]
75. [ndārúm ^w á páfūrò]	I've been bitten on the neck.	75. [h]

In the following minimal pairs the fricatives under investigation are in contrastive distribution with other consonantal phonemes in the Karanga system:

- /-sará/, "stay here" vs. /-zará/, "be full".
- /-díwá/, "be feared" vs. /-kíwá/, "be eaten".
- /-fárá/, "rejoice" vs. /-párá/, "scratch"
- /-sótá/, "annoy" vs. /-rótá/, "dream" (v.i.)
- /-zará/, "give birth" vs. /-zará/, "be full".
- /-xòrà/, "earn, be paid" vs. /-kòrà/, "be fat".
- /yud̥zi/, "hair" vs. /rud̥zi/, "race" or "type"(n.cl.ll)
- /hāngá/, "guinea fowl" vs. /bāngá/, "knife".
- /m^wezá/, "mouse run above ground" vs. /m^wená/, "hole in ground"

5.3.10.2. The Search for Acoustic Cues.

Many interesting studies of the fricatives of a number of languages have been made with a view to isolating those acoustic features which most characterise these sounds. Only a few are summarised below for purposes of comparison with those of Karanga.

Lieberman's (1957)²⁰ paper is a summary of the work in acoustic phonetics done at the Haskins Laboratories in the previous twenty years. He cites three acoustic cues found to be important for the perception of fricative sounds.

First, the location, on the frequency scale, of the fricative noise, and in particular the lower frequency limit of this noise, was found to be an overriding cue for distinguishing American English [s] from [ʃ]. As we shall see below, this is also a crucial feature distinguishing [ʒ] from [s], and each one of them from [ʃ]. This variable, however, does not seem to contribute much to the perception of other, less intense fricatives such as [f] and [θ].

The second cue reported by Liberman has to do with the formant transitions of vowels adjacent to the fricatives. These transitions, particularly F2 and F3 transitions, give significant information regarding the place of articulation of the sound in question.

Finally, the American English group [s, ʃ, z, ʒ]

20. Liberman, A.M. (1957) "Some Results of Research on Speech Perception", J.A.S.A. 29,1, pp.117-123.

was found to be distinguished from [f, θ, v, ð] by the general intensity level.

Stevens' (1960)²¹ paper is a detailed spectro-graphic analysis of what may be loosely called "cardinal consonants". As the title of the paper suggests, the fricatives studied were not from any one particular language. He described the fricatives with a view to specifying (i) the frequencies of the lower and upper limit of energy present;
(ii) the presence or absence of formant-like concentrations of energy;
(iii) the overall relative intensity of the sounds.

The sounds investigated, like the plosives of Halle, Hughes and Radley (1957)²², fall into three groups:

Front: [f, θ, ð]

Central: [s, ʃ, ʒ]

Back: [x, χ, h]

He found the following acoustic features to be common to these classes of fricatives:

Front: (a) Long spectrum, with little patterning of peaks of power.

(b) Relative intensity low.

Central: (a) Short spectrum, with main region of energy at higher frequency than in the other two groups.

(b) Relative intensity high.

21. Stevens, P. (1960) "Spectra of Fricative Noise in Human Speech" Language and Speech, 3.1.

22. Halle, Hughes and Radley (1957) op.cit.

Back: (a) Spectrum medium length.

(b) Exhibit a formant-like patterning of energy.

(c) Relative intensity is intermediate between the other groups.

Strevens makes two general points. The first is that the shape of the vocal tract has a direct bearing on the resulting sound - stressing the important interdependence of physiological and acoustic events. Change in this shape is probably the chief cause of perceptible differences in sound quality.

The second point made is that the pressure of expiratory air used in speech is operated on by two variables:

- (a) affective variations of pressure: the pressure level for shouting and other animated speech is higher than for normal speech.
- (b) phonetic impedance of pressure: the air-stream above the larynx has two modes of flow, aperiodic and periodic, equivalent to breath and voice, respectively, with the former always at a higher rate.

In either mode of flow, the air may be free or impeded to a greater or lesser extent. Both the extent and duration of such obstruction vary from sound to sound.

The stops studied above have momentary but complete obstruction while fricatives have lower phonetic impedance of longer duration.

By taking general phonetic considerations into account, and by extrapolating his findings in the analysis of the voiceless fricatives above, Strevens made some

general remarks about voiced fricatives and other sound types, having a noise component. The voiced counterparts of the fricatives above, [β , v, ∂ , z, ʒ , j, ʒ , ʃ], were characterised by Strevens as consisting of two components: a component of hiss, and a component of 'vocal tone'. The acoustic characteristics of the hiss obviously correspond to those of the voiceless fricatives in most respects. The vocal tone corresponds to the 'voice-bar'.

The main difference between the two types of fricative is that in the articulation of a voiced fricative, for a given air-pressure the air-flow is less than for a corresponding voiceless fricative. The acoustic intensity of the hiss component of voiced fricatives is inherently less than that of the voiceless.

Strevens makes one important final point,

"The statements are an approximation and a normalization, and not absolute or final descriptions. Palatographic studies show that there are often large variations of shape and orifice and even of place of constriction, within the speech of an individual²³ as well as between speakers It is clear that there are available to the speaker, compensatory processes which enable him to produce an acceptable quality of voiceless fricative using quite a variety of different articulatory postures."

The possibility of there being articulatory (and, consequently, acoustic) differences, of greater or lesser extent, between the realizations of what is judged by native speakers and competent phoneticians as the "same", acceptable sound, is clearly recognised by Strevens.

23. Compare the two Palatograms for each laminal sound studied in this thesis.

This has important implications for the spectrograms, palatograms and mingograms analysed in this thesis. Their absolute, etic values may differ from those of the sounds produced by another speaker of Karanga: but such differences do not violate the system, the emic value of the material which permits inter-personal and inter-group communication within the dialect.

5.3.10.3. The labio-dental fricatives /f,ɸ/.

(a) Pronunciation.

As the diagram below shows, the articulatory stricture for both the fricatives is the same.

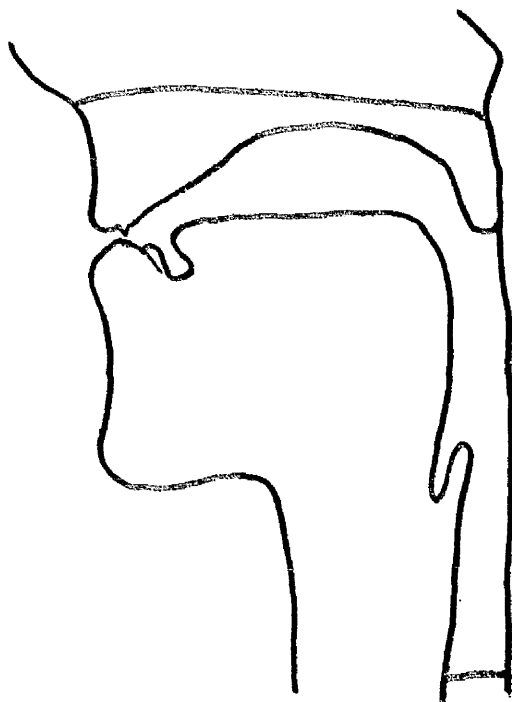


Diagram 13. /f,ɸ/.

There is a narrow degree of approximation between the upper incisors and the inside surface of the lower lip.

The soft palate is raised, cutting off the nasal chamber. For the voiceless fricative [f], the vocal cords are wide apart, thereby allowing a big volume of aperiodically vibrating pulmonic air to flow into the pharynx-mouth cavity; Karanga [y̥] is a breathy-voiced fricative. During its pronunciation, the ligamental vocal cords vibrate as for plain voice but the arytenoid cartilages are wide apart, producing both voice and breath. Throughout, the tongue assumes the position of the adjacent vowel.

For both sounds, air issues from the speech tract fricatively because of the narrow labio-dental stricture. The air has a median but slit mode of egress.

(b) Acoustic Features.

Both [f] and [y̥] are typical fricatives in that they have no formant structure on their random-looking spectrograms 63 and 64. The voiceless fricative has greater intensity compared to its breathy-voiced counterpart. For both sounds, the fricative noise is spread on the spectrograms from about 1.00 KHz, where it is very weak, to the very highest frequencies. Overall the intensity is rather low, as can be seen, for example, by comparing the amplitude sections of [s], 65 and of [f] on section 63. Although a number of utterances were used, only a very low amplitude section could be obtained for the breathy-voiced [y̥] - section 64, as in [pánemáyud̥z̥ɪ].

The vowel preceding both fricatives on these spectrograms is the same, [a]. It has a terminal F2

transition pointing towards a very low locus, at about 700Hz, as does the initial F2 transitions of the vowel following each fricative in the respective utterances. This low formant transition locus is characteristic of the Front series of consonant articulations generally.

Apart from the intensity difference between them, these two fricatives are also distinguished by the presence of a voice-bar for [y] at about 250 Hz, missing on the spectrogram of [f]. This voice-bar has striations with more energy next to the vowels on either side of the fricative, than at the centre of it. The vowel preceding [f] is about 14 centi-seconds in duration, while that preceding [y] is about 18 centi-seconds. This is an approximation: it is easier to see the beginning of [f] than that of [y] because of the break in voicing between it and the preceding vowel.

5.3.10.4. The dental fricatives /s,ʒ/.

Pronunciation.

During the pronunciation of these two fricatives, the soft palate is raised to exclude nasal resonance. Pulmonic egressive air is impeded by the articulatory narrow stricture between the tip and blade of the tongue and the tooth ridge. For [s] the Palatograms (i) a and b show that the sides of the tongue make firm contact with the left and right alveolar zones. The surface of the tongue tip and blade is grooved and the air forces its

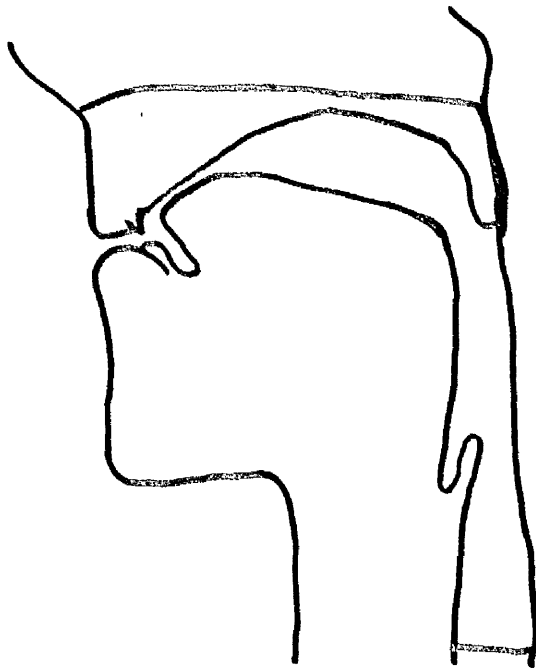


Diagram 14. /s̺, z̺/.

way along the groove and through the interstices of the upper incisors with a sharp hissing sound. Lip position depends on that of the adjacent vowel. For example, the lips will be spread in pronouncing the example for Palatogram (i) a [s̺imbá], whereas they will be open wide for the example (i) b [pásá].

The purpose of having two Palatograms is clearly demonstrated by comparing the area of side contact for [s̺] before a high front vowel and before an open one. There is far more side contact especially between the front of the tongue and the sides of the hard palate in [s̺imbá] than in [pásá]. In other words, there is less interference from the vowel in the latter example than in the former, although both varieties of [s̺] are classified as members of the same distinctive /s/ phoneme.

Much the same features may be observed on Palatograms (ii) a and b [z̥imbá] and [pàz̥à], respectively, for the breathy-voiced dental fricative, except that this sound seems to have even wider contact between the sides of the front of the tongue and the area of the hard palate between the second molar line and fourth molar line than its voiceless counterpart before a front vowel.

(b) Acoustic Features.

As spectrograms 65. [sára muma'sure] and 66. [pave' namazue'a] show, the noise component of these two fricatives is concentrated in the high frequency range, above 3.5 KHz. The voiceless sound has greater overall intensity than the breathy-voiced one.

The vowel preceding both fricatives, [a], and [u], the vowel following, both have a very low locus F2 transition, at about 800 Hz, consummate with the articulations being dental or denti-alveolar. The noise component is in the high frequencies mainly because they are grooved articulations, whereas [θ, ð], which are produced in the same area, are slit, for example.

There is also a difference in the duration of the vowel preceding the voiceless [s], 14 centi-seconds and that preceding [z̥], about 19 centi-seconds. The other difference between these two sounds is that [z̥] has a voice-bar with low intensity and wide-apart striations, whereas [s] does not have a voice-bar at all.

While the overall profiles of the amplitude sections 65 and 66 are similar, amplitude section 66, for

[ʒ] has no energy near the middle frequency region, at points corresponding to those at which [s] has its lowest level of energy. This is to be expected since an open glottis allows more air to pass upwards from the lungs than when the vocal cords are vibrating.

Compared to [f,y], [s,ʒ] have most of their energy in the high frequency region, whereas the former show no preferential concentration of such energy on the frequency axis: it is evenly spread all over, more or less.

5.3.10.5. The 'labialised' alveolar fricatives /ʃ,ʒ/.

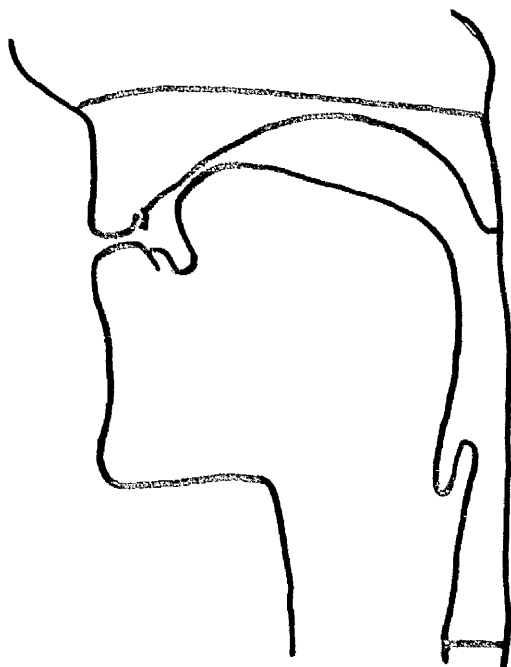


Diagram 15. /ʃ,ʒ/.

(a) Pronunciation

For both sounds the soft palate is raised to cut off nasal resonance. For [ɕ], the glottis is wide open, allowing aperiodically vibrating air to flow into the pharynx-mouth cavity; whereas for [ʒ] the glottis produces simultaneous breath and voice (see Ladefoged (1971), pp. 6, 8).

The narrow articulatory stricture for these sounds is made between the blade of the tongue and the alveolar ridge, as Palatograms (iii) a and b and (iv) a and b show. There are two major differences between these palatograms and those for [s, z]. First, the passage for the pulmonic air is wider than that for [s, z]. Secondly, the closure between the sides of the tongue and the left and right alveolar zones extends as far forward in the mouth as the alveolar ridge for [ɕ, ʒ], and farther for the dental fricatives.

When palatograms (iii) a and b are compared it is found that, as in their dental counterparts, the area of contact between the sides of the tongue and the left and right alveolar zones is wider when [ɕ] is followed by a high vowel, as in [ɕímbo], than when it is followed by an open vowel, as in [ɕàvə]. Overall, the contact between the sides of the tongue and the alveolar zones seems to extend farther forward in the mouth in [ɕí-] than in [ɕà-], though remaining alveolar. This also holds for palatograms (iv) a and b, for [ʒ].

The third major difference between the alveolar pair and the dental one is the lip position during the pronunciation of these sounds. The alveolar sounds are

called "labialised" because during their pronunciation the lips are closely rounded, irrespective of which vowel(s) is adjacent, whereas [s,ʒ] are accompanied by anticipatory lip-rounding only when the following vowel is [o] or [u]. Lip-rounding for [ʃ,ʒ] gives these sounds that peculiar friction which has earned them the label "whistling fricatives".

(b) Acoustic Features.

The overall amplitude section profiles for the two sounds are similar, with energy peaks and minima at corresponding points along the frequency axis - see amplitude sections 69 and 70, for [ʃ] and [ʒ] respectively.

There is an energy peak at about 3.5 KHz, which declines progressively thereafter, more precipitately for [ʒ] than for [ʃ]. The former sound has no energy along the frequency axis at a point corresponding to that with the lowest energy concentration for [ʃ].

On spectrograms 69. [təm^wá mǎʒùʒú] and 70. [áʒúvǎ mǎsǒ] the lower limit of noise for the two sounds on the frequency scale is at 2.5 KHz, compared to 3.5 KHz for [s,ʒ], and, as we shall see below, at 2 KHz for [ʃ,ʒ].

The vowel [a], preceding each of the two sounds in their respective utterances, has F2 and F3 frequency transitions having loci at slightly higher frequencies than for [s,ʒ] above, that is, between 800 Hz and 1 KHz. The following vowel has a clear initial F2 transition pointing towards the same frequency area.

The two fricatives, [ʃ,ʒ] are distinguished one from the other by the presence of a voice-bar on the

spectrogram for [ɣ], but missing on that for [ɣ̥].

The vowel [a] has a duration of 20 centi-seconds before [ɣ̥], compared to 16 centi-seconds before [ɣ].

5.3.10.6. The alveolar lateral fricatives /ɣ̥, ɣ̥̥/.

(a) Pronunciation.

To produce these sounds, the soft palate is raised, cutting off nasal resonance. For [ɣ̥], the vocal cords are held wide apart, allowing aperiodically vibrating pulmonic air to flow into the pharynx-mouth cavity. During the pronunciation of [ɣ̥̥] on the other hand, there is simultaneous voice and breath, as for all breathy-voiced consonants.

The articulatory strictures for these fricatives are made in two different parts of the oral cavity, as palatograms (vii) [ɣ̥íwá] and (viii) [ɣ̥̥íwá] show.²⁴ The forward closure is made between the blade of the tongue and alveolar ridge. The second stricture is a narrowing between the right side of the back of the tongue and that part of the right alveolar ridge behind the fourth molar. The alveolar closure is maintained, so that pulmonic air has a lateral egress out of the speech tract.

If we decide that the closure in the mouth, which makes these sounds lateral, is more important, then we

24. This is the only pair of fricatives for which one set of palatograms (before a front vowel) has been made. An inspection of all the other examples where a palatogram was made in the environment of an open vowel will show that there is a fairly general pattern and the features displayed thereon may be extrapolated onto the lateral fricatives.

will have to call them alveolar sounds. But if, on the other hand, the stricture which produces friction (thereby making the sounds fricative) is chosen, then [ɗ, ʒ] will be called back lateral fricatives. Our classificatory criteria, outlined in the second chapter, favour the former solution: a closure ranks higher than a narrowing; and there are more median oral strictures in speech than lateral, i.e. using the criterion of articulatory potential.

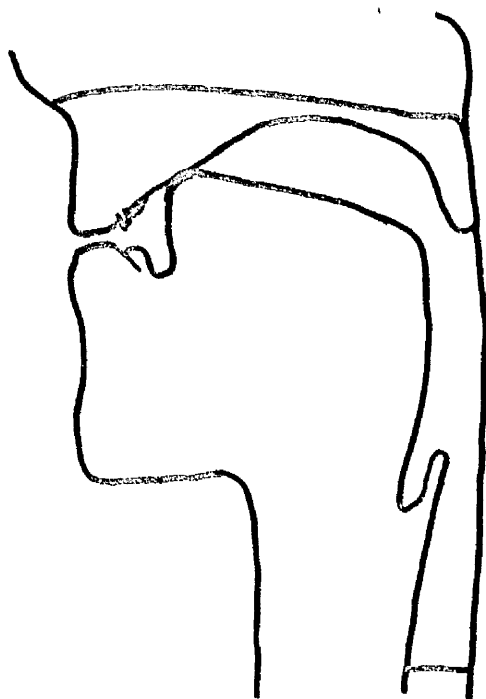


Diagram 16. /ɗ, ʒ/.

(b) Acoustic Features.

The problem outlined above is not resolved on the acoustic level since the features observed on the 67 and 68. [tāvapakùgá] and [ùnónàná kùdá] respectively, spectrograms are conflicting. As seen on spectrograms 67 and 68. [tāvapakùgá] and [ùnónàná kùdá] respectively,

[d, ɖ], like the back set of fricatives [x, ɣ, ʁ], have a clear formant structure²⁵. They each have four formants with centre frequencies at 250 Hz, 2.25 KHz, 3.75 KHz and 5 KHz respectively. Especially for [ɖ], these formants correspond to four vowel-type peaks of energy on the amplitude sections 67 and 68.

But like other central articulations, [d, ɖ] are characterised by vowel formant transitions with a fairly high locus. Both [u] and [a], the vowels preceding and following the fricatives respectively, have F2 frequency transitions pointing towards a locus of about 1.9 KHz. So the two sounds have acoustic features which correspond to the two areas of the mouth in which the strictures which produce them are made.

The sounds are distinguished one from the other by the following features:

- (i) [d], the voiceless fricative, has greater intensity than the breathy-voiced [ɖ] - compare the amplitude sections 67 and 68;
- (ii) [ɖ] has a low frequency voice-bar with wide-apart striations, whereas [d] has none;
- (iii) the vowel [u] has a duration of 20 centi-seconds before [ɖ] compared to about 12 centi-seconds before [d].

5.3.10.7. The palato-alveolar fricatives, /ʃ, ʒ/.

(a) Pronunciation.

The soft palate is raised to cut off nasal resonance.

25. See P. Strevens (1960) op.cit.

For [ʃ] the vocal cords are held wide apart, thereby allowing a large volume of aperiodically vibrating air into the supra-glottal cavities. For [ʒ], the vocal cords are positioned for producing breathy-voice.

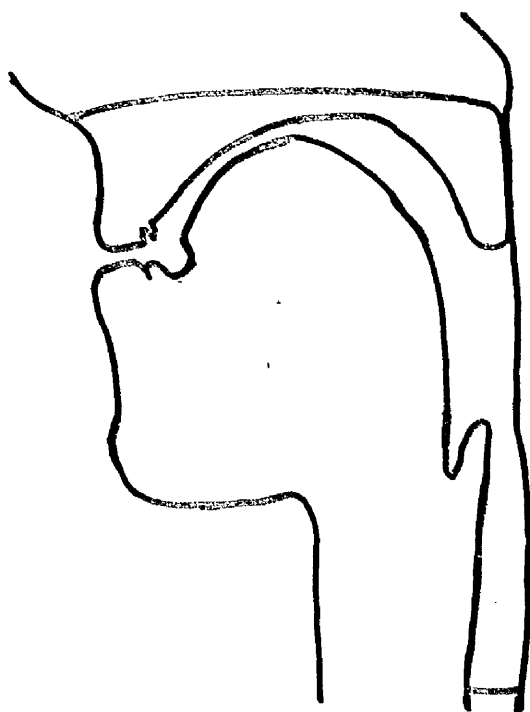


Diagram 17. /ʃʒ/.

As the palatograms (v) a and b, and (vi) a and b show, the sides of the tongue come into firm contact with the left and right alveolar zones. As in the previous examples, in the examples a, the fricative is followed by a close vowel, whereas the examples b have a low vowel which does not interfere with the pronunciation of the consonant. Taking the voiceless [ʃ] as an example, it is clear that there are marked differences between the [ʃ] in [ʃúmbà] and that in [ʃàmbà]. For

example, the air passage seems to be wider in example (v) b than in a. The area of contact between the sides of the front of the tongue and the left and right alveolar zones is wider in [ʃambà] than in the other example. This is perhaps due to the fact that [a], the following vowel is neither back nor front: though open, this vowel has tongue raising in about the same area as the palato-alveolar fricatives [ʃ, ʒ]. There is, therefore, some interference.

However, both examples with a high and a low vowel have side contact which extends from behind the fourth molar line to as far forward in the mouth as the canine line, which is not as far forward as for the other fricatives above. And compared to them, that is [s, ʒ, ʃ, ʒ], [ʃ, ʒ] have a much wider groove for the air from the lungs to pass through, causing a less sharp hiss than in those fricatives. The lips are rounded, but not as closely as for [ʃ, ʒ].

(b) Acoustic Features.

The amplitude sections of [ʃ] and [ʒ] have a similar overall profile, with peaks of energy at corresponding points along the frequency axis. On average, amplitude section 71, for [ʃ], has higher energy peaks than that of the breathy-voiced [ʒ]. And the latter has no energy at all at the point corresponding to that where [ʃ] has its lowest energy concentration.

The relative intensity of these fricatives is lower than that of either [s, ʒ], or [ʃ, ʒ]. The lower limit of the noise component on the frequency scale for

[ʃ,ʒ] is at about 2 KHz.

On spectrograms 71 [z̥aári māʃurá] and 72. [ndàwánà m̥z̥um^{wí}], the vowel [a], preceding both fricatives in these utterances, has a fairly high terminal F2 transition locus at about 1.9 KHz, as does [u] after the fricatives.

The two sounds differ from each other in that [ʒ] has a voice-bar whereas [ʃ] does not. The vowel [a] has a duration of about 18 centi-seconds before [ʒ], and of 14 centi-seconds before [ʃ].

[ʒ].3.10.8. The velar fricatives /x, ɣ/.

(a) Pronunciation.

The fricatives [x] and [ɣ] are unvoiced and breathy-voiced respectively. In the case of [x], aperiodically vibrating air is allowed to flow into the supra-glottal cavities through the open glottis. For [ɣ] the vocal cords are positioned for breathy-voice as described by Ladefoged (1971, pp.6,8).

The soft palate being raised to cut off nasal resonance (as shown in the diagram below), pulmonic air for both [x, ɣ] flows into the pharynx-mouth cavity producing friction as it forces its way through the narrow stricture between the back of the tongue and the soft palate. The position of the lips depends on that of the adjacent, especially the following, vowel. Pulmonic air has a median mode of egress.

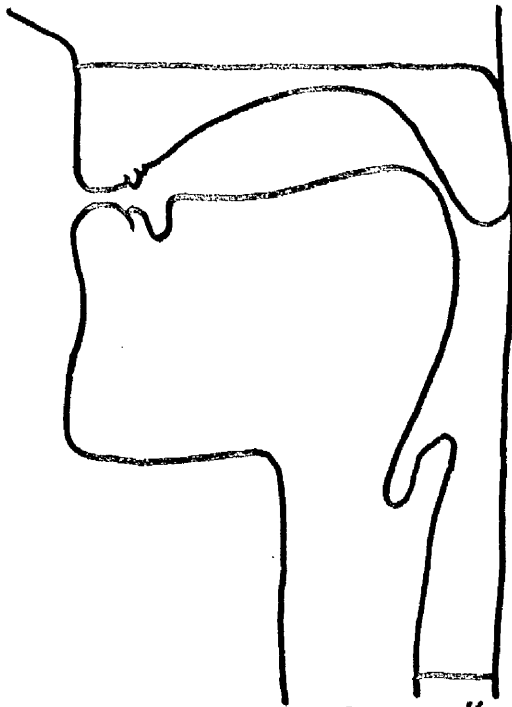


Diagram 18. /x,χ/.

(b) Acoustic Features.

The amplitude sections of both fricatives have lower level energy than all the other fricatives in the Karanga system except the labio-dentals [f,ɸ]. Their overall amplitude profiles are similar, with vowel-type peaks of energy at corresponding points along the frequency axis. And the intensity difference between [x] and [χ] is minimal, with the former having slightly higher peaks. The voiceless [x] has a low level energy concentration in the very high frequency region where the breathy-voiced [χ] has none - compare the amplitude sections 73 and 74.

Spectrograms 73 [t̪axórà n̪ási] and 74 [nd̪inódá'χówà] show that both fricatives, like [d,ɓ] above, have

a clear formant structure. Each one has a low frequency F1 (725 Hz), a rather high F2 (3.5 KHz), and F3 (4.5 KHz).

Apart from having a formant structure like [d, b], there are no outstanding similarities between these two pairs of sounds. For one, the centre frequencies of corresponding formants are different, those of [x, ɣ] being higher than those of [d, b]. Secondly, the terminal F2 and F3 frequency transitions of [a] before [x] and [ɣ] point towards a common locus at about 1 KHz. This, according to Ladefoged (1975, p.178), among others, is a defining characteristic of all velar articulations.

Between themselves, [x] and [ɣ] differ in that the latter has a voice-bar for about the first quarter of its duration, whereas [x] has no striations at all. The vowel preceding [x] has a duration of about 15 centi-seconds, whereas that preceding [ɣ] has a duration of about 18 centi-seconds.

5.3.10.9. The glottal fricative /ħ/.

(a) Pronunciation.

The Karanga glottal fricative is the breathy-voiced consonant par excellence. With the arytenoid cartilages open and the ligamental vocal cords vibrating, the larynx produces a complex of both voice and aperiodic friction.

The soft palate being raised to exclude nasal

resonance, this lung air is channelled into the pharynx-mouth cavity. The shape and size of these supra-glottal cavities depends on the following vowel, which means that there are as many h -quality sounds as there are vowels in the system (see chapter 3 above). In that chapter a number of similarities were observed between the amplitude sections of $[\text{h}]$ and that of the associated vowel in the so-called hV -sequences. This is why some writers²⁶ prefer to regard the glottal fricative as the voiceless, or (in our case) the less voiced onset of the ensuing vowel.

Pulmonic air is expelled from the speech tract at high pressure, producing friction in all the supra-glottal cavities, except the nasal.

(b) Acoustic Features.

Amplitude section 10 shows that $[\text{h}]$ has a regular arrangement of vowel-type peaks of energy. On this particular section, where the fricative precedes the vowel $[\text{u}]$ in the syllable $[\text{hu}]$, only the first two formants can be seen.

The amplitude section also shows that the fricative has its highest concentration of energy in the very low frequencies, like the first two formants of $[\text{u}]$, up to about 700Hz. This energy decays rather rapidly above this frequency, more so on amplitude section 10 than on 75.

Spectrogram 75, shows that the vowel preceding

26. Doke, C.M. (1931 p.92) and Gimson, A.C. (1962,p.191) op.cit.

[f̥] in [ndàrúm^{wá} páf̥urò] has very rapid terminal formant transitions. But there are no transitions between the fricative and the ensuing vowel [u], confirming our articulatory claim that during [f̥], the supraglottal cavities are shaped as for the following vowel. The fricative itself has striations which are wider apart than those of the vowels on either side.

5.3.10.10. Conclusion.

At the outset, and following Strevens (1960)²⁷, the fricatives of Karanga were divided into three classes based on articulatory criteria, namely Front, Central and Back. The following are some of the more outstanding acoustic features which characterise these classes:

Articulatory Features

Acoustic Features

Front:

[f, y]	Very low loci for the F2 and F3
[s, z]	vowel transitions, in the frequency range 250 Hz - 800 Hz.

Central:

[ʃ, ʒ]	Very high loci for F2 and F3 vowel
[t, d, ʈ, ɖ]	transitions, from 1.5 KHz to 2 KHz.

27. Strevens, P. (1960) op.cit.

Back:

[x, ɣ, ɦ]

- (i) Rather low frequency loci, for F2 and F3 vowel formant transitions, between 500 Hz and 1 KHz.
- (ii) Formant-like concentrations of energy.

Grooved articulations:

[s, z; ʃ, ʒ]

The noise component is concentrated in the high frequencies on the spectrum with that of [s, z] having the highest lower limit at 3.5KHz, that of [ʃ, ʒ] the lowest at 2KHz.

Laterals:

[ɬ, ɮ]

Formant-like concentrations of energy, like the Back series.

The members of the Front class are distinguished among themselves by the low intensity which characterises the labio-dentals [f, v], as opposed to that of [s, z], which is much higher. Members of the Central series are distinguished among themselves by [ɬ, ɮ] having a formant structure which [s, z; ʃ, ʒ] have not got; and the latter by the lower limit of the noise component of [ʃ, ʒ] being higher than that of [ʃ, ʒ].

The Back series is made up of [x, ɣ], which have F2 and F3 vowel transitions pointing towards a common locus, and [ɦ], which does not have vowel formant transitions which are very clear since it is itself very similar to neighbouring vowels.

5.3.10.11. The Affricates of Karanga.

Introduction.

As I said in the section on plosives above, the release of the mouth closure during the pronunciation of a plosive sound is radical, that is, clear-cut. If, however, the two organs involved in this pronunciation part slowly, we perceive homorganic friction, not plosion. Such friction is aperiodic if there is no voicing during the closure period; and periodic if there is voice or (in the case of Karanga) breathy-voice, during the closure period.

Sounds produced in this way, that is, with plosive-like closures followed by homorganic friction, are called "affricates". Gimson (1962, p.171)¹ says

"The term 'affricate' denotes a concept which is primarily of phonetic importance. Any plosive, whose release stage is performed in such a way that considerable friction occurs approximately at the point where the plosive stop is made, may be called 'affricative'. The friction present in an affricate is of shorter duration than that which characterises a fricative proper."

Later in the same chapter Gimson (1962, p.172)²⁸ says that it is possible to regard English affricates either as single phonemic entities or as sequences of two phonemes:

"If /t , d , tr, dr/ are treated as sequences of two phonemes, the realization of the second element will differ according to whether it occurs in the same syllable or morpheme as the stop...".

28. Gimson, A.C. (1962) op.cit.

I favour a unitary phonemic interpretation of affricates for Karanga (and Central Shona generally) where consonantal phonemes do not occur in syllable final positions.

In terms of the definition of 'affricate' given by Gimson above, Karanga has a set of eight affricates:

pf, by; ts, dz; tʃ, dʒ/. In the section on plosives, /p, b/ have been classified as bilabials, and /f, v/ as labio-dentals in the section on fricatives.

The first two items in the list of Karanga affricates are therefore not made up of homorganic elements. But they nonetheless qualify for this classification since the two elements are not required to be strictly homorganic: Gimson says "considerable friction occurs approximately at the point where the plosive stop is made..."

There are a number of distinguishing features peculiar to affricative sounds. For example, affricates combine in sequence some of the articulatory features of stops (e.g. momentary but complete closure of the air-stream) with other features characteristic of fricatives (e.g. partial obstruction of the air-stream).

In the following study, use will be made of spectrograms of utterances in which the affricates are surrounded by the same vowels thus: aAu, where 'A' stands for the affricate in question, and [a] and [u] are the vowels preceding and following the affricate respectively:

76. [màpfúmo' áwándà], "there are now many spears".

77. [màbyúra' húkú], "you(pl.) have singed a fowl".

78. [ndínódà tsúngà], "I want 'vegetables'".

79. [áǎ dʒùngú], "(s)he has stolen pea-nuts".
80. [zínòndàtsúkà], "I'm now light-skinned".
81. [ndìnóǎ dʒúkú], "I want the red one" (adj. cl.5).
82. [ndámupà tʃúmí], "I've given him/her ten shillings".
83. [ndákámáǎʒurú], "I've eaten ants".

In addition, Palatograms were made for all these affricates except two, /pf, by/. As for all the other sounds produced in the palatal and alveolar areas, two Palatograms were made for each affricate, one using an utterance in which the affricate preceded a high vowel and the other in which the vowel was low. The palatograms so obtained were numbered (xviii) a and b through (xxiii) a and

5.3.10.12. The labio-dental affricates /pf, by/.

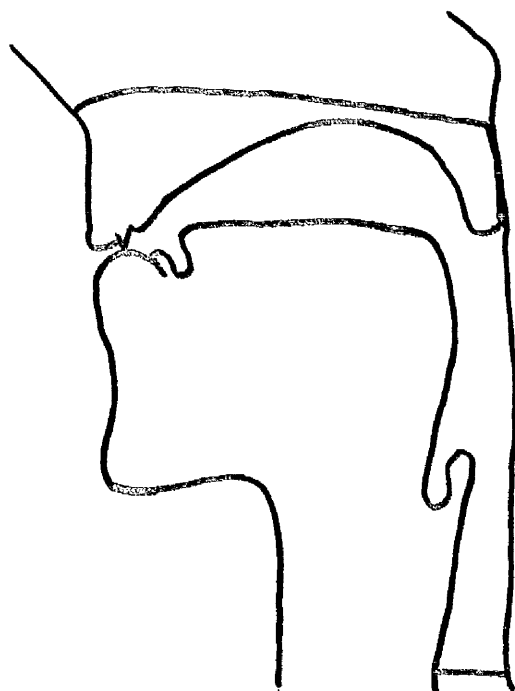


Diagram 19. /pf, by/.

(a) Pronunciation.

To produce these affricates, the soft palate is raised to cut off the nasal chamber. The obstacle to the pulmonic egressive air-stream is a complete closure

formed by the inner surface of the lower lip being pushed up against the upper incisors. Over and above this stricture the upper lip comes very close to, sometimes touches, the edge of the lower lip.

The tongue is not directly involved in the pronunciation of these affricates, and so its position will be that appropriate for the following vowel. Lung air therefore builds up pressure behind the labio-dental closure until this pressure exceeds that of atmospheric air. The closure is then released slowly, causing audible friction.

During both the stop and fricative stage, the vocal folds are wide apart for [pf], allowing a strong current of aperiodically vibrating air to flow into the supraglottal cavities. For [by], on the other hand, the vocal folds are positioned for producing breathy-voice.

(b) Acoustic Features.

On spectrogram 76. [màpfumó áwándà], there seems to be a distinct noise "burst" following the release of the closure stage, apart from the friction of [f]. The fact that such a noise "burst" is not seen on spectrogram 77. [màbyúrañúkú], may be accounted for by

the pressure behind the closure being stronger for the voiceless [pf] than for the breathy-voiced [b_y].

The most outstanding acoustic features are the F2 and F3 terminal transitions of the vowel [a] before these affricates in their respective utterances. The downward movement of the transitions of the F2 and F3 of the vowels [a] is characteristic of transitions from the steady-state of the formants of vowels to a labial or dental or labio-dental consonant articulation. In this case these formant transitions point towards loci at about 200 Hz and 1.5 KHz respectively.

One of the interesting features on all the spectrograms of affricates in this study is the behaviour of the first formant of the vowel [a]: it has a transition bending downwards before all the affricates irrespective of their place of articulation. This seems to be some corroboration of Delattre's (1968)²⁹ observation that whereas F2 and F3 transitions give information about the place of articulation of a consonantal sound, F1 transitions give information about the manner of articulation of such sounds. The downward movement of F1 of the vowel preceding an affricate seems to be an underlining feature of this class of sounds.

The other outstanding feature on these spectrograms is the diffuseness of the noise component of the affricates: it is spread from the very low to the highest frequencies, being rather low in intensity.

29. Delattre, P. (1968) op.cit.

The two affricates differ from one another by [b_y] having a voice-bar with far-apart striations while [pf] does not have any voice-bar. What seem to be striations at about 250Hz on the spectrogram of [pf] made me record more utterances and analyse them using the Speech Segmentator to find out if this sound has voice in intervocalic positions: those are not striations and [pf] has no voice in those positions.

5.3.10.13. The denti-alveolar affricates /ts, d_z/.

(a) Pronunciation.

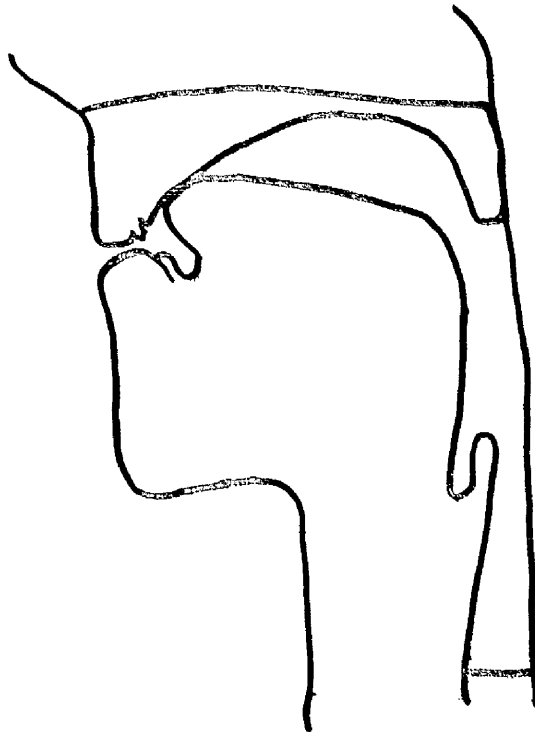


Diagram 20 . /ts, d_z/

The soft palate is raised to cut off nasal resonance. And, as Palatograms (xviii) a and b, and (xix) a and b show, there is firm contact between the sides of the tongue and the left and right alveolar zones. This contact starts at the extreme back of the palate, behind the fourth molar line, and extends as far forward in the mouth as the lateral incisor line for the examples a, that is, [tsímè] and [d̥zìmbá], where the affricates are followed by the front high vowel. In the examples b, on the other hand, that is, [tsámà] and [d̥zámà], where these affricates precede a low vowel, the contact stops just short of the lateral incisor line.

This difference between the two sets of Palatograms is accounted for by the overlapping between the consonantal and the vowel articulations. But this particular instance of interference is minor compared to that along the sides of the oral cavity. Here there are marked differences in the width of the area of contact between the sides of the tongue and the left and right alveolar zones in the context of the two vowel articulations. These differences are most marked in the areas on either side of zones 6 and 7, where the contact has a width averaging 14 mm and 21 mm respectively for examples a; and 12 mm and 13 mm in zones 6 and 7 respectively for the examples b.

For both examples, however, pulmonic egressive air builds up pressure behind the denti-alveolar closure, which pressure is higher for the voiceless [ts] than for the breathy-voiced [d̥z]. The closure is then released slowly in the centre but is maintained on the sides.

5.3.10.14. The 'labialised' alveolar affricates /tɕ, dʒ/.

(a) Pronunciation.

The soft palate is raised to cut off nasal resonance during the pronunciation of these affricates as shown in the diagram below.

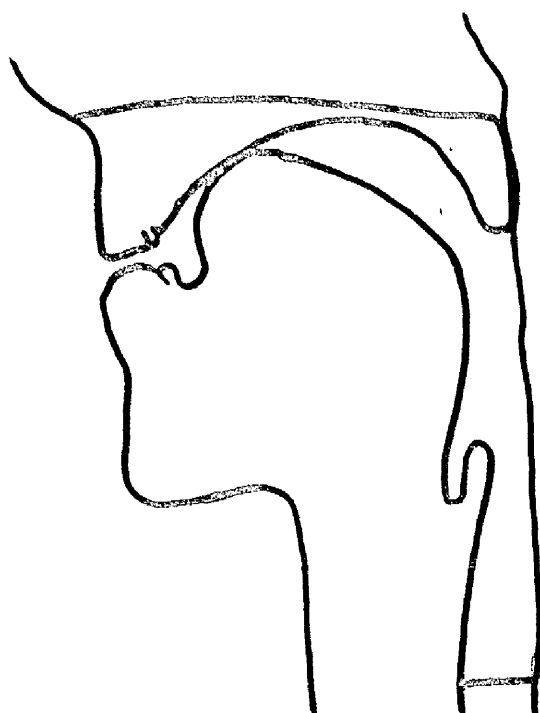


Diagram 21. /tɕ, dʒ/.

And as the Palatograms (xx) a and b, [ìtɕá] and [màtɕá] respectively, and (xxi) a and b, [ìdʒá] and [màdʒa] respectively, show, the air from the lungs is obstructed by a firm closure between the sides of the tongue and the left and right alveolar zones, continuous with one between the tip and blade of the tongue and the alveolar ridge.

In both cases of example a, where the affricates are preceded by the front high vowel, the central area contact comes very close to the lateral incisor line. On the Palatograms of examples b on the other hand, the closure between the alveolar ridge and the tip and blade of the tongue is less advanced, exceeding the canine line only by a few millimetres. The extent of the interference of the front high vowel is again underlined. This is even more noticeable when the width of the side closure with a front vowel preceding is compared with that with an open vowel preceding the affricates.

This side contact is much wider when the high vowel is adjacent to the affricates than when an open vowel is involved. Overall, it seems that the contact between the sides of the tongue and the left and right alveolar zones is slightly wider on the left than on the right, irrespective of which vowel is adjacent to the consonantal articulation - and this applies to most Palatograms in this study.

In addition to the strictures described in the preceding paragraphs, these two affricates require close lip-rounding and a certain degree of lip-protrusion, which give the sounds their distinctive "whistling" noise, also a feature of the fricatives [ɬ, ʂ].

During the closure period the vocal cords are wide apart, allowing aperiodically vibrating air to flow into the supraglottal cavities for [tɬ], whereas they are positioned for producing breathy-voice for [dʂ]. In either case, pulmonic air builds up pressure behind the

oral closure until this pressure exceeds that of atmospheric air. The central closure is then released slowly, causing alveolar friction as the compressed air escapes from the speech tract.

(b) Acoustic Features

On spectrograms 80 [zínòndàtsúkà] and 81. [ndìnóda d̥ukú], the first formant of the vowel [a] before [t̥, d̥] in their respective utterances, has a falling terminal transition pointing towards a very low locus, as in all the other examples. Although these affricates are alveolar articulations, the second formants of the vowels flanking them do not point towards as high a locus as one has come to associate with such sounds. This formant transition points towards a locus of about 1KHz. This may be due to the close lip-rounding in the pronunciation of these affricates.

Also on both spectrograms, the noise component of the affricates is concentrated in the frequency range from 2 KHz upwards, though there is some energy below that limit. Although there are some striations on those portions of the affricate [d̥] which are adjacent to the vowels surrounding it, they have a very low level of energy; and the middle portion is not much different to the whole of the closure and fricative portions of [t̥], which have no voice.

5.3.10.15. The palato-alveolar affricates /tʃ, dʒ/.

(a) Pronunciation.

Nasal resonance is excluded by the raised soft palate, as shown in the diagram below.

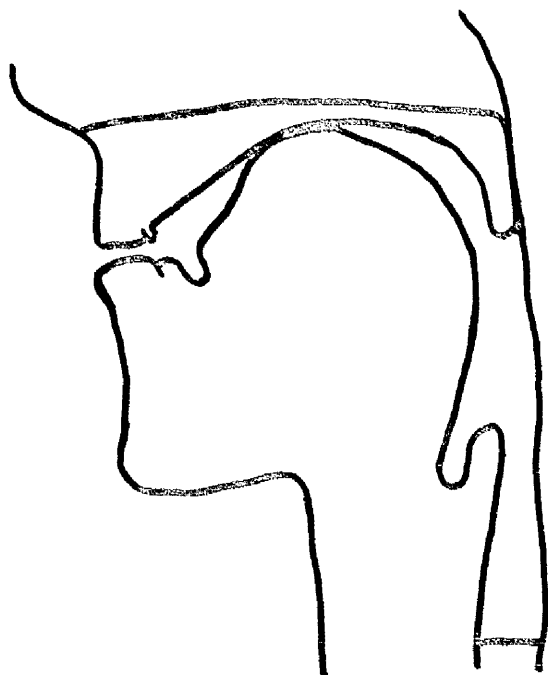


Diagram 22. /tʃ, dʒ/.

And, as can be seen on the Palatogram (xxii) a and b, [tʃipó] and [tʃápá] respectively, and (xxiii) a and b, [mùdʒíngá] and [dʒává] respectively, there is firm contact between the sides of the tongue and the left and right alveolar zones. This closure extends from behind the fourth molar and extends as far forward in the mouth as the canine line. Central area contact is between the palato-alveolar arch and the blade of the tongue. This central contact covers a wider area in the

examples a than the examples b. On Palatogram (xxii) a for example, this closure covers the area of the palate between the second molar line and the canine line. On Palatogram (xxii) b on the other hand, it lies between the first molar line and the canine line.

That is not the only difference between these two sets of Palatograms: the contact between the sides of the tongue and the alveolar zones is wider when the affricates are adjacent to high vowels than when they are not.

In addition to these strictures, there is a certain amount of lip-rounding and protrusion which occurs whether or not the affricates are followed by rounded vowels. But these lip modifications are minor compared to those characteristic of labialised affricates proper.

During the closure period, the vocal cords are wide apart for [tʃ], so that aperiodically vibrating air enters the mouth-pharynx cavity, whereas they are positioned for producing breathy-voice for [dʒ]. In both cases, lung air is compressed behind the palato-alveolar closure as for a plosive. Release of this closure is partial, that is, only the central area contact is slowly removed, while that along the sides is maintained. This produces palato-alveolar friction as the air escapes through the mid line of the oral cavity .

(b) Acoustic Features

On Spectrograms 82. [ndámúpà tʃúmí] and 83. [ndáǵá máǵúurú] the first formant of the vowel [a] has an initial transition pointing downwards, as in the other examples discussed above. The second and third formants of that vowel, on the other hand, have transitions pointing upwards to higher loci than in all preceding examples. The second formant has a locus of about 2.5 KHz, while the third one points towards one at about 3 KHz. The second vowel [u] has much the same loci for its corresponding formants.

The noise component of these affricates is concentrated in the frequency region above 1.5 KHz; although there is some energy below this limit, it is weak.

The breathy-voiced affricate has a voice-bar with wide-apart and low-intensity striations. The unvoiced [tʃ] has no voice-bar, although there are some striations at the beginning of its hold stage before the vocal cords have attained their steady-state for the affricate.

5.3.10.16. The labio-velar "affricates" /px, bɣ/.

(a) Pronunciation.

In addition to the regular affricates described above, Karanga has two sounds which cannot be called affricates in terms of the definition given by Gimson (1962, pp.171-172) and quoted at the beginning of this section on affricates. In my pronunciation of these sounds [px, bɣ] the plosives have a separate pronunciation

as they do in isolation, that is, closure, silence, release and plosion. The pronunciation of the fricative commences only after these stages have been completed: this assertion is supported by my Kinaesthetic sensations and evidence from the Speech Segmentator.

The plosion of the stop is weakened by the narrow stricture formed between the back of the tongue and the soft palate for the fricatives [-x] and [-ɣ], pronounced immediately after the plosive is released. These fricatives reduce the force of the air that was compressed behind the bilabial closure. The vocal cords are wide apart for the voiceless [px], while they are positioned for producing breathy-voice for [bɣ]. Throughout this pronunciation, there is a velic closure, so that nasal resonance is excluded.

(b) Acoustic Features.

These two articulations share, with the regular affricates discussed above, the feature of a falling terminal F1 transition of the preceding vowel. This can be seen on Spectrograms 84. [kùràpxà] "to be cured", and 85. [gòndóbɣe], "a ram".

As before all labial articulations, the vowels next to these sounds have F2 and F3 transitions pointing towards low loci at about 600 Hz and 1 KHz respectively. But the vowels [a] and [e], which immediately follow [x-, ɣ-] in their respective utterances, have different transition patterns altogether. The second and third formants of [a] after [x] point towards a common locus

at about 1.8 KHz, while those of [e] after the breathy-voiced [ɣ] are not too clear.

The release of the bilabial closure in both utterances is followed by a noise burst, with more energy after the voiceless than after the breathy-voiced plosive. This noise component is spread from the lowest to the highest frequencies, unlike that of the fricatives, which has its lower limit at about 1 KHz, and is arranged in formant-like patterns.

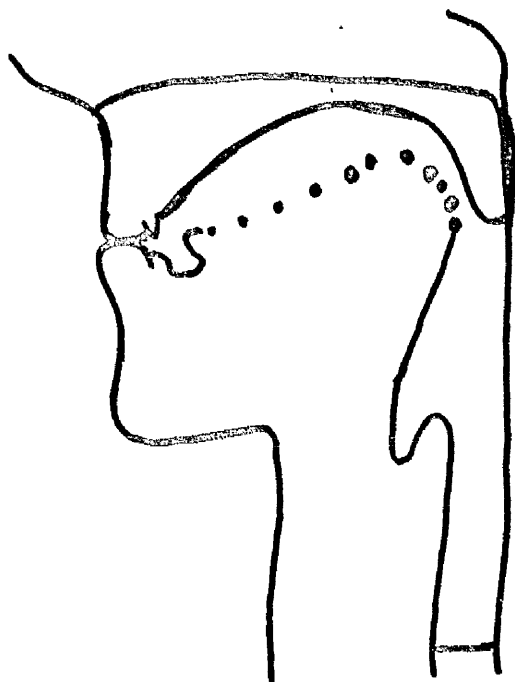


Diagram 23. /px, bɣ/.

Finally, to illustrate the use of the breathy-voiced

"affricate" above, consider the following tongue twister

coined by speakers of other dialects to parody Karanga

[bɣowa bɣuja bɣatakajanika paruware bɣaxenganiɬiwa

nemadzibɣa egondobɣe bɣukadzoka bɣofabɣabɣa / saka tobɣira
 ɬɔɣowa ɔɣuja ɔɣatakajanika paruware ɔɣaxenganiɬiwa

nemadzibɣa egondobɣe bɣukadzoka bɣofabɣabɣa / saka tobɣira

hédú tsígíí], "that mushroom which we left in the sun to dry has been mixed with ram mucus. As a result it has become rancid. So we have to eat sugar instead." This sentence also shows that the "affricate" [bʒ] may occur in the syllable before any of the five vowel phonemes of Karanga.

5.3.11.0. The nasal phonemes.

Introduction.

Karanga (and Central Shona as a whole) has a set of four nasal consonants with what we have called "plain voice": during their pronunciation, the vocal cords are made to vibrate periodically throughout their entire length. These articulations are produced at four places in the speech tract; the bilabial, alveolar, palatal and velar: /m,n,ɲ,ŋ/. They occur contrastively with other phonemes in the dialect:

/-márá/, "scratch" vs. /-sárá/, "stay behind" vs.
/-nárá/, "be shy".

/ɲòmbè/ "cattle" vs. /ròmbè/, "vagabond".

/màrà/, "home area" vs. /-nàrà/, "plaster" (v.t.).

In their pronunciation of these consonants the speech organs are lax, Kinaesthetically speaking, compared to the tension which characterises the pronunciation of the corresponding, breathy-voiced set of nasals:

/ṁ,ṇ,ṅ,ṁ/. The phonemic status of some of these consonants may not be too transparent if we insist on finding minimal pairs in which the plain-voiced nasal contrasts with the breathy-voiced. But each one of

these breathy-voiced nasals occurs contrastively with other phonemes in the system:

/mándá/, "animal fat" vs. /ṁándá/, "fork in branch of tree".

/màrí/, "money" vs. /ṁàrí/, name of sub-dialect of Karanga and of ruling clan.

This gives us two bilabial nasal phonemes /m,ṁ/. Each of these phonemes contrasts independently with other Karanga phonemes:

/mándá/, "animal fat" vs. /-vándá/, "hide" (vɪ) vs. /ṁándá/, "fork in branch of tree". And also, /màrí/ "money" vs. /ṁàrí/ name of Karanga sub-dialect. vs. /tʃàrí/, "shawl".

The alveolar nasal with plain voice contrasts with its breathy-voiced counterpart in the following minimal pairs:

/-nàkà/, "be nice" or "sweet" vs. /ṅàkà/, "inheritance";
/nòkó/, "from your end" vs. /ṅòkó/, "goat droppings";
/-nàv̀à/, "wilt" or "wither" vs. /ṅàv̀à/, "fibre woven bag".

As in the case of the bilabials above, these alveolar nasals enter into independent paradigms with other consonants in the system. Thus:

/-nàv̀à/ and /ṅàv̀à/ each contrast with /-ràv̀à/, "read" or "count", while /nòkó/ and /ṅòkó/ each contrast with /ʃòkó/, "monkey".

Although the other two sets of nasals provide no minimal pairs as between a plain-voiced one versus its breathy-voiced counterpart, they nonetheless contrast with other consonantal phonemes in the system.

The palatal plain-voiced and breathy-voiced nasals occur in the following minimal pairs:

/-núrá/, "drown" vs. /-kúrá/, "grow up"

/-nàpà/, "overdo" vs. /-tsàpà/, "fast" (v.i.).

/mènú/, "alive" (adj. cl.9/10) vs. /mènú/,
invitation to baby to ride on mother's back.

/pási/, "down" or "on the ground" vs. /pási/,
direction downwards.

/pxàrà/, "idiotic chap" vs. /pxàrà/, "crush" (v.t.).

The velar pair also occurs in contrastive positions with other phonemes in the system:

/ŋàngà/, "herbalist" vs. /wàngà/, "scar" (n.5).

/-ŋàrà/, "complain bitterly and demonstratively,"
vs. /ràrà/, "dirt", (n.5).

/-rùrà/, "be vicious" vs. /-rùrà/, "be tamed by
punishment".

/-ŋòrà/, "become spiritually unclean" vs. /-xòrà/,
"earn" or "be paid".

In view of the examples cited above, we may posit eight nasal phonemes for Karanga, using two parameters : place of articulation and phonation type:

<u>Phonation type</u>	<u>Place of articulation</u>			
	(a)	(b)	(c)	(d)
	<u>bilabial</u>	<u>alveolar</u>	<u>palatal</u>	<u>velar</u>
(a) plain-voiced:	m	n	ɲ	ŋ
(b) breathy-voiced:	ṁ	ṇ	ɲ̤	ŋ̤

As in the previous sections, we can divide these articulations into three sub-groups, calling the bilabials

the Front, the alveolars and palatals the Central, and the velars the Back series. The alveolars are grouped together with the palatals because they are the so-called 'laminal' articulations, i.e. the tongue makes a closure with the roof of the mouth in pronouncing them. Later we shall see that as a result of this, these two sets share some acoustic features not found in the other nasals.

5.3.11.1. Acoustic Studies of Nasals.

The one most obvious feature characterising the spectra of nasal consonants, and one which many researchers have remarked on, is their vowel-type but low-intensity formant structure (see Fujimura [1962]³⁰; Fant [1970]³¹; Delattre [1968]³² and Ladefoged [1975]³³).

Fant (1970)² says that a voiced nasal occlusive is characterised by a spectrum in which F2 is weak or missing. And, more importantly, a formant at approximately 250Hz generally extending into the adjacent vocalic segment "dominates" the spectrum.

Ladefoged (1975 ,p.178)³³ notes that /m,n,n/ have

30. Fujimura, O. (1962) "Analysis of Nasal Consonants." Journal of the Acoustical Society, 34, pp.1865-1875.

31. Fant, G. (1970) op.cit.

32. Delattre, P. (1968) op.cit.

33. Ladefoged, P. (1975) op.cit.

vowel-type albeit low-intensity, formants; that these formants lie in particular frequency locations that depend on the characteristic resonances of the nasal cavities. He speaks of "nasal cavities" because each nasal sound has a particular nasal cavity-cum-oral cavity combination.

Consequently, the centre frequencies of the formants of these nasals differ depending on the location of the oral occlusion for a particular nasal.

In specifying the nasals of Karanga, attention will be focused on the following acoustic features:

- (a) the formant structure, i.e., the centre frequencies of F1, and of F2 and F3, if present;
- (b) the formant transitions of adjacent vowels, in particular F2 and F3 transitions;
- (c) how the breathy-voiced vs. plain-voiced dichotomy is reflected on (i) spectrograms;
 - (ii) amplitude sections;
 - (iii) Mingograms.

5.3.11.2. The bilabial nasals /m,m/.

(a) Pronunciation

To produce these nasal sounds (as the diagram below shows) the lips come together, forming an airtight closure. The soft palate is lowered, coupling the oro-pharyngeal chambers with the nasal.³⁴ Since these are

³⁴. For all nasal consonantal articulations, the oral cavity is semi-active, whereas the pharyngeal and nasal cavities are active, since air from the lungs enters, but does not pass through the oral cavity, while it passes through the other two cavities.

not laminal articulations, the tongue assumes its position for the adjacent vowel.

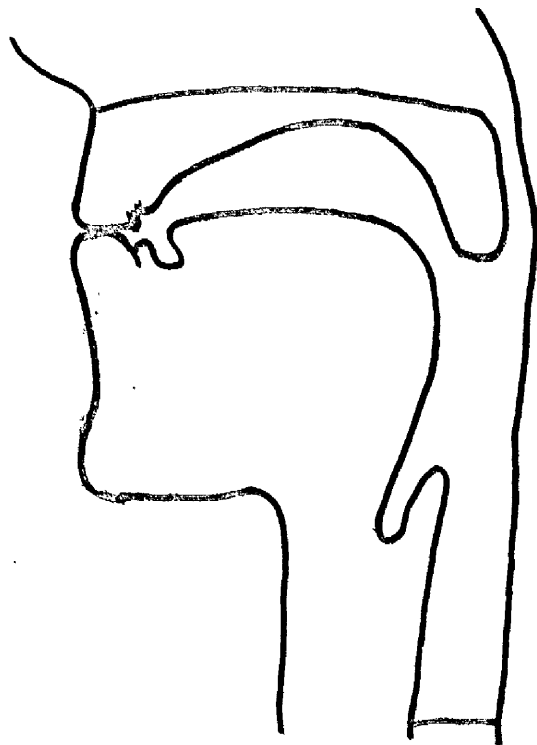


Diagram 24. /m, m̠/.

For the plain-voiced [m], the vocal cords vibrate periodically whereas they produce murmur or breathy-voice for [m̠]. The auditory impression conveyed by these two sounds reflects the muscular tension during their production. The sound [m̠], like all breathy-voiced sounds in the system, gives a rough, guttural auditory impression, whereas [m], the Kinaesthetically lax sound, is higher and less throaty.

(b) Acoustic Features.

The following observations are based on spectrograms of utterances in which the sounds [m, m̠] are in syllable initial positions preceding different vowel

qualities:

- 86. [tʃú mí], "ten shillings",
- 87. [mándá], "animal fat"
- 88. [mándá], "fork in branch of tree"
- 91. [mé nù], "in your (pl)..."
- 92. [mè nù], "alive" (adj. cl.9/10).

On spectrogram 86. [tʃú mí], the bilabial nasal has a clear formant structure; and its formants may be described as having low intensity only in comparison with those of the vowels surrounding it in the utterance. The following formants are clearly visible: F1 at about 250 Hz; F2 at about 1.25KHz; F3 at 2.25 KHz and F4 at 3 KHz. There is what may be termed an "anti-resonance" (Fant, 1960, p.143)³⁵ from about 3.5KHz to 5KHz: there is no energy at all. Above that frequency range, energy is concentrated at 5.5KHz, 5.75KHz and 6.25KHz.

The vowel following [m] in this utterance, i.e. [i], has a fairly steep initial F2 transition pointing towards a very low locus at about 500Hz. The third formant also seems to have an abrupt transition pointing downwards.

We have two sets of spectrograms in which the plain-voiced nasal and its breathy-voiced counterpart are in comparable phonetic environments: 87 and 88, and 91 and 92. Taking the latter two examples first, a number of differences between [m] and [m̥] can be observed:

- (i) longer duration for [m̥] (about 18 centi-seconds) than for [m] (about 12 centi-seconds).
- (ii) the striations on the formants of [m̥] are wider apart

35. Fant, G. (1960) op.cit.

than those of [m].

(iii) the plain-voiced [m] has more energy in the higher frequencies above the third formant than the breathy-voiced [ṁ].

The second set of examples, 87 and 88, are, in fact, spectrograms-cum-amplitude sections of these two sounds. The aim in making these sections was to find out how the phonatory differences between [m] and [ṁ] are reflected thereon. It is easier to see, with the naked eye, differences in the sizes of amplitude peaks than the minute differences in the spacing of striations on spectrograms. The amplitude peaks of both nasals correspond to the number of formants visible on the spectrograms, but those of the plain-voiced nasal are higher - a tracing of one of the amplitude sections was made on transparent paper and superimposed on the other section for comparison.

Further, Mingograms³⁶ were made of the utterances 4. [mándá] "animal fat", 5. [ṁándá] "fork in branch of tree". 6. [mìndà] "plots of land", 7. [ṁìndò] "darkness" (of night). Oscillograms for the first two examples show that [m] has greater voice amplitude than [ṁ], and that this low voice amplitude for [ṁ] continues throughout the duration of the syllable. Not only are the oscillations (corresponding to striations on spectrograms) for [ṁ] of lower amplitude

36. Care was taken to choose examples in which [m] and [ṁ] precede the "same" vowels in syllables bearing the same tone. In order to economise on space, the Mingograms of the other nasals, i.e. alveolar, palatal and velar, have been omitted. The features observed on them are the same as those of the bilabials.

but they are also farther apart than those of [m]. This means that the vocal folds are more ideally positioned for voice production during the pronunciation of [m] than that of [ṁ]. This has some bearing on the pitch of the syllables of which these nasals are the initial segment, as can be seen by comparing the pitch tracings on Mingograms 4 and 5. In fact [ṁ] has virtually no pitch tracing on Ming. 5, and although it has on Ming. 7, this is lower than that of its plain-voiced counterpart. This is why breathy-voiced consonants are labelled "depressors" in Shona tonal phonology (see Chapter 7).

5.3.11.3. The alveolar nasals, /n,ṇ/.

(a) Pronunciation.

As the Palatograms (xiii) a and b show, there is firm contact between the sides of the blade and front of the tongue on the one hand, and the left and right alveolar zones on the other. There is not much difference in the width of this contact as between Palatogram (xiii) a [póná], and b [páná] along the sides. There is, however, a difference in the size of the central area closure between the tip and blade of the tongue, and the alveolar ridge: that of the example a lies between the canine line and half-way between the first and second molar lines; whereas that of the example b is narrower, lying between the first molar line and the canine line.

Throughout the pronunciation of the alveolar nasals the soft palate is in a lowered position, coupling the oro-pharyngeal cavity with the nasal. The vocal folds produce plain voice for [n], and breathy-voice for [ɲ], as pulmonic air flows between them. The lips assume their position for the adjacent vowel since they are not directly involved in the pronunciation of [n,ɲ] - see Diagram 25 below.

(b) Acoustic Features

The acoustic correlates of the phonatory contrasts between plain-voiced [n] and breathy-voiced [ɲ] may be observed on spectrograms 89. [nàkà], and 90. [ɲàkà], "be nice" and "inheritance" respectively,

On each of these spectrograms, three vowel-type formants are clearly visible. Their centre frequencies are at 250Hz, 1.75KHz and 2.75KHz on both spectrograms. But the second and third formants of the plain-voiced [n] have a much darker outline than the corresponding formants of the breathy-voiced [ɲ].

The other difference between these two nasals is that the striations of [n] are closer together than those of [ɲ]. The accompanying amplitude sections for these sounds are also informative: they each have three peaks of energy corresponding to the three formants visible on the spectrograms. When the tracing (on transparent paper) of the amplitude section of [n] was superimposed on the section of [ɲ], the former was found to have higher peaks of energy than the latter. The voice amplitude difference between these two nasals is greater than that between the

bilabials, and the initial F2 transition of the vowel [a] on each spectrogram has a high locus at about 2.3KHz.

5.3.11.4. The palatal nasals, /ɲ,ɲ/.

(a) Pronunciation.

The Palatograms (xiv) a and b show that on average, the palatal nasal has the widest area of contact between the sides of the tongue and the left and right alveolar zones of all the nasals, irrespective of what the adjacent vowel is. On Palatogram (xiv) a [ɲimò], central area contact covers the central zones 4 and 5, and small fractions of zones 3 and 6: on example b, on the other hand this area is smaller, covering only central zone 4 and a small portion of 3, since there is less overlapping from the vowel [a] of [ɲàmà].

On the whole, the differences between the palatal nasals [ɲ,ɲ] and their alveolar counterparts [n,n], auditory and acoustic, may be attributed more to the extent of the lateral closures than to the central area closure being more or less advanced in the mouth. This comes out clearly when the Palatograms of these two subsets of the central nasals are compared.

During the pronunciation of the palatal nasals, the soft palate remains lowered, allowing pulmonic air to have a nasal egress. The lips assume their position for the following vowel. And the vocal cords produce plain voice for [ɲ] and breathy-voice for [ɲ̤].

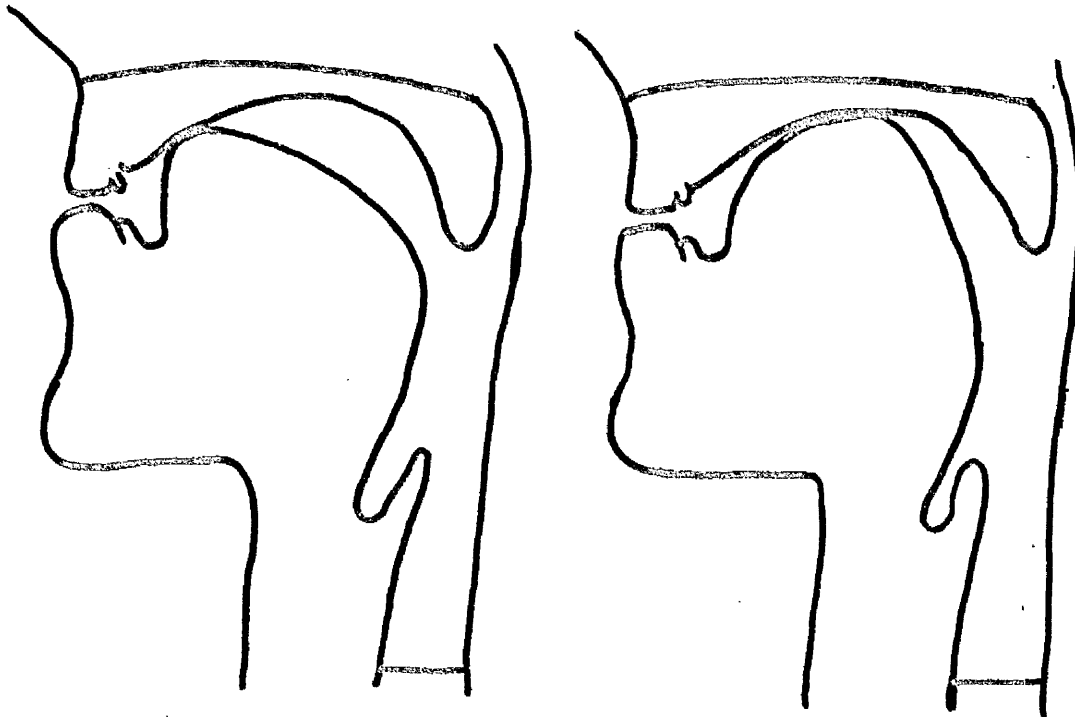


Diagram 25 . /n,ɲ/. Diagram 26 /p,ɸ/.

(b) Acoustic Features.

On spectrograms 91. [mé'nú] and 92. [mè'nú] the two palatal nasals are followed by the "same" vowel in a high tone syllable. Both sounds have a formant pattern, although that of the plain-voiced [p] has bolder outlines than that of its breathy-voiced counterpart.

The most outstanding feature on these spectrograms is the initial F2 and F3 transitions of the vowel [u]. These formant transitions have a common locus at about 2.3KHz. The coming together of these two formant transitions, as we have seen in previous sections, is a characteristic feature of velar consonants. But on spectrogram 93. [ní'mò], where the plain-voiced [p] precedes the front high vowel [i], these formants do not point

towards a common locus.

These observations lead to the conclusion that the palatal nasals [ɲ,ɲ̥] are affected by the quality of the following vowel: there is some assimilation between the vowel and the consonant articulation. These two sounds have retracted allophones before back vowels, and fronted ones before front vowels. This means that it may not be too easy to distinguish [ɲ,ɲ̥] from the velars [ŋ,ŋ̥] in the context of back vowels solely on acoustic (especially formant transition) evidence, although this is easily done using auditory and Kinaesthetic sensations.

5.3.11.5. The velar nasals, /ŋ,ŋ̥/

(a) Pronunciation.

To produce these two sounds, the back of the tongue makes a closure with the velum, as in the pronunciation of the velar plosives [k,g]. The difference between these two pairs of sounds in this respect is that whereas there is a velic closure for the plosives, there is a nasal passage for pulmonic air during the pronunciation of [ŋ,ŋ̥]. The position of the lips depends on the following vowel in an utterance.

As with the palatal nasals, pulmonic egressive air has two modes of flow past the glottis: the vocal folds vibrate periodically, producing plain voice for [ŋ], but are positioned for producing breathy-voice for [ŋ̥].

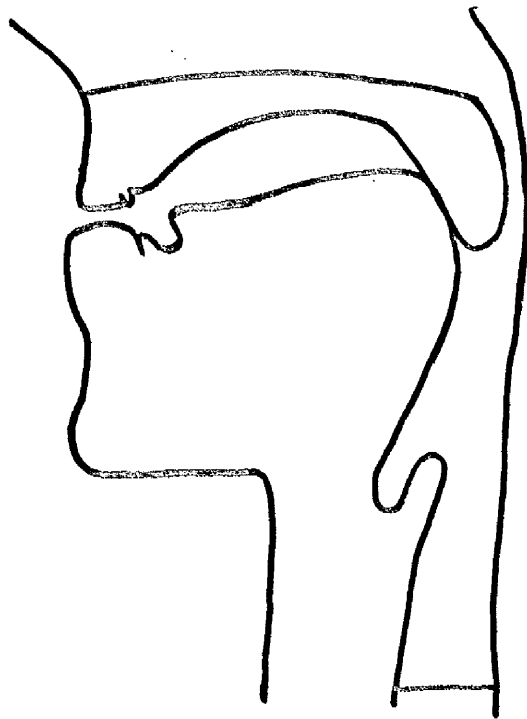


Diagram 27. /ŋ, ɲ/.

(b) Acoustic Features.

Two spectrograms, 94. [ɲùrà] and 95. [ɲòrà], each with an amplitude section of the velar nasal in question, were made. There are some obvious common features between these spectrograms. For example each nasal has a formant structure comprising three low-intensity formants: F1 at 350 Hz; F2 at 1.6KHz; and F3 at 2.75KHz.

These formants correspond to three peaks of energy on the amplitude sections. Equally, there are some differences between these sounds. For example, the voice-bar of the plain-voiced [ŋ] has more energy than that of the breathy-voiced [ɲ]; the amplitude peaks of the former sound are higher than those of the latter nasal. On spectrogram 96 [ɲéɲéná] the initial F2 and F3 transitions of the vowel [e] in the first syllable point

towards some fairly high loci at about 1.6KHz and 2KHz respectively - although, this consonant [ŋ] being a velar, I had expected the two transitions to have a common locus, as they do next to [k, g, x, ʁ].

Conclusion

Taking the observations made in this section into account, the following seem to be some of the defining acoustic correlates of the articulatory criteria used in the general introduction to sub-classify the nasals of Karanga.

Articulatory Features:

Acoustic Correlates:

Front:	m } ɱ }	Low frequency loci for both F2 and F3 vowel transitions - not exceeding 700Hz.
Central:	n, ɲ } ɳ, ɳ } (before front vowels)	Loci of F2 and F3 of the vowel transitions as high as 2 KHz and 2.3 KHz respectively.
Back:	ŋ, ɳ } ɰ, ɰ } (before back vowels)	Loci of F2 and F3 vowel transitions are intermediate between the Front and Central series.

The retracted varieties of [ɳ, ɰ] have a common locus for F2 and F3 transitions.

On the other hand, the phonatory dichotomy plain voiced versus breathy-voiced has the following acoustic correlates:

<u>Plain-voiced</u>	<u>Breathy-voiced</u>
[m, n, ɲ, ŋ]	[<u>m</u> , <u>n</u> , <u>ɲ</u> , <u>ŋ</u>]
(i) Striations close together	(i) Striations farther apart than for plain-voiced counterparts.
(ii) Vowel following has striations close together	(ii) Part of vowel next to nasal has wide apart striations.
(iii) Does not seem to lower syllable pitch.	(iii) Causes gliding pitch from relative low to relative high.
(iv) Amplitude sections have higher peaks than for breathy-voiced counterparts.	(iv) Amplitude sections have lower energy peaks.

In his Standard Shona Dictionary Hannan³⁷ uses an orthographic convention whereby the distinction between breathy-voiced sounds on the one hand, and plain-voiced ones on the other, is indicated by the presence or absence of the letter "h" after the nasal in a prevocalic position (nasals being taken as examples; this applies to all other sounds in the system). For example, the words used in this study are represented in that Dictionary as follows:-

<u>Phonemic Representation</u>	<u>Orthographic Representation</u>
/mándá/	manda, animal fat
/ṁándá/	mhanda, fork in branch of tree
/mìndà/	minda, plots of land

37. Hannan, M., S.J. (1974) op.cit.

<u>Phonemic Representation</u>	<u>Orthographic Representation</u>
/m̀ɪndò/	mhindo, darkness (of night)
/-nàkà/	naka, be sweet, nice
/m̀ɪndò/	nhaka, inheritance
/-nàkà/	menyu, in your (pl.)...
/nàkà/	*mhenyu, alive adj.9/10.
/mɛ́nú/	n'ura, remove meat from bone
/mɛ́nú/	using teeth.
/-ɲùrà/	*n'ora, become spiritually
/-ɲòrà/	unclean.

This new orthography is a commendable attempt to incorporate an important phonetic (and phonological) feature in the writing system. My only regret is that this has not been extended to all instances of breathy-voiced syllable onsets. For example, in my pronunciation the palatal nasal represented by "ny-" in the first asterisked example above is breathy-voiced and should therefore be written "nyh". Similarly the velar nasal represented by "n'" in the second asterisked example should also have an "h", thus "n'h..".

5.3.12.0 The Naso-Oral Phonemes.

Introduction.

The phenomenon which writers on Shona phonetics and phonology have described by the term "prenasalization" is one whereby the breathy-voiced as well as the plain-voiced plosive consonants, the breathy-voiced fricatives (except /ɸ/) and the affricate /ɕ/, combine with the homorganic or semi-homorganic, appropriately voiced,

nasal. Such combinations have been treated, and I think correctly, as functional units, i.e. each as one complex consonantal onset to the syllable. The voiceless consonants do not combine with nasals in this way, as can be seen on the following list of so-called "prenasalised" consonant complexes: /mb, m̥b, m̥y, nd, n̥d, n̥z, n̥ʒ, ndʒ, ŋg, ŋg and m̥bʏ/.

As I have pointed out in the first two chapters, terms such as "velarization", "palatalization", "labialization" etc. are usually used in the phonetic literature to describe the addition of a secondary articulatory feature to a primary articulatory feature: for example, the environmentally conditioned rounding of the lips for an otherwise lip-spread or neutral fricative such as [ʒ] in [zúʒa] "sun" or "day", where lip-rounding is a modifying feature caused by the following lip-rounded vowel [u].

It is because I use these terms in this way that I am unhappy with the term "prenasalization". According to the criteria outlined in the second chapter, a closure ranks higher than an opening of whatever degree if both strictures are made simultaneously during the pronunciation of a particular sound. In these so-called "pre-nasalised" complexes, when the soft palate is lowered for the nasal sound, there is simultaneously a closure in the mouth. This is why nasals are called bilabial, alveolar, palatal or velar nasals: because that is where the articulatory oral closure is made. During such strictures, the nasal cavity is active (p.99-101 above), i.e. pulmonic air enters the cavity and leaves the speech tract that way; whereas

the oral cavity is semi-active.

To the extent that there is some oral resonance during the pronunciation of the nasal consonant, it may be phonetically plausible to speak of an oralised nasal - since the contribution of the nasal cavity to the quality of the resultant sound is superior to that of the oral. But the consonant complexes described as "prenasalizations" are an entirely different matter: because whereas the strictures for the nasal consonant are simultaneous, those for the nasal as a whole, and those for the oral sound with which the nasal combines, are in sequence.

So it does not seem reasonable, in the latter case, to talk of an oral sound which has been prenasalised. The resulting consonantal complex is best described as a combination of a nasal and an oral sound, and not a prenasalised oral, which would imply that the nasal is merely a modification of the oral sound. In this study it has been found, on the contrary, that the nasal consonant in these complexes is almost invariably longer in duration, has more voice amplitude, higher pitch and greater intensity than the oral sound which it precedes in these combinations... henceforth to be called naso-oral sounds, plain- or breathy-voiced. There would seem, therefore, to be no acoustic or articulatory justification for describing the bilabial plain-voiced naso-oral sound [mb], for example, as a prenasalised oral plosive.

Research for this section consisted in choosing lexical items and grammatical constructions in which the naso-oral sounds occur in syllable initial position. Data from such examples was obtained using the Electro-Aerometer

such that information was recorded on Mingograms on a number of dimensions: mouth tracing; nasal tracing; larynx activity shown on oscillogram; pitch; and intensity both on the linear and logarithmic scale.

These Mingograms are particularly useful in the description of these sounds because the relative durations of the elements of these complex sounds can be easily measured, and any vowel nasalization noted. And since the Electro-Aerometer was calibrated, the pitch at which the segments were produced can also be measured fairly accurately.

The mingograms used in this description are of the following words and phrases, bearing the same number as the actual mingograms:

- | | |
|---|--|
| 8. [sím [́] bá], "power, strength" | 9. [sím [́] bà], "genet" |
| 10. [m [́] yúrá], "water, rain" | 11. [ndíró], "plate"(n.9/10). |
| 12. [ndíró], "it's the one" | 13. [n [́] zùngú], "peanuts" |
| 14. [n [́] zimbò], "place"(n.9/10) | 16. [nd [́] z [́] íwá], "dove" |
| 15. [n [́] zírè], "trickster" | 17. [ngùrùwè], "pig" |
| 18. [ngùpání], "child's intestinal disorder". | |
| 19. [zímbyéndé], "very cowardly person". | |

In addition, a number of lexical items and grammatical constructions were also recorded and data obtained using the sound spectrograph. These are also useful because they provide acoustic evidence for what is felt Kinaesthetically to be a sequence rather than a simultaneity of naso-oral consonant articulations:

- | | |
|---|------------------------------------|
| 97. [sím [́] bá], "power, strength". | 98. [sím [́] bà], "genet" |
| 99. [ndìnóda m [́] yúrá], "I need water" | |
| 100. [ndíró], "plate". | 101. [ndíró], "it's the one". |

102. [kètánzúmà], "choose those (cattle) without horns".
 103. [ndànáńzá], "I have licked...".
 104. [règánkò dzákó], "stop your greediness".
 105. [ndípè màndzúg^wé], "give me ground crickets".
 106. [ngùrúwè], "pig". 107. [ngùbání], "child's intestinal disorder".
 108. [mb^yírem^yíré], "meal eaten dry".

Palatograms were made for those naso-oral sounds pronounced in the alveolar and post-alveolar regions of the mouth. Where the dichotomy applies, it was not found necessary to make separate Palatograms for the plain-voiced as well as for its breathy-voiced counterpart - since this is a distinction that applies to the phonatory rather than the articulatory strictures. Only the alveolar [nd] and [nd̤] are affected:

- (xii) a [ndípè], "give me..."; b [mándá], "animal fat".
 (xv) a [nzímbe], "sugar cane" b [wànzà], "increase in quantity".

(xvi) a [nzímbo], "place"(n.9/10)

b [nzà], ideo. of carrying on shoulder.

(xvii) a [nzó], "greediness" b [nzàngà], "grain store"

(xxiv) a [ndzìwá], "dove" b [màndzà], "applause!"

5.3.12.1. The Labial naso-oral phonemes.

(a) The plain-voiced bilabial naso-oral phoneme, /mb/.

(1) Pronunciation

Both segments involved in this complex phoneme have a bilabial closure. During the pronunciation of

the first segment, the soft palate remains lowered, allowing pulmonic air to flow out of the speech tract through the nasal cavity. Some of this air enters the oral cavity prior to being diverted to the nasal, as during the pronunciation of the plain bilabial nasal [m].

The nasal cavity is then cut off by a velic closure for the second segment [b]. This results in all the lung air flowing into the oral cavity. With the bilabial closure being maintained, this air is compressed until its pressure is higher than that of atmospheric air. Throughout, the vocal cords are held close enough together to vibrate periodically when air from the lungs passes between them. But since the capacity of the supra-glottal cavities is limited, air can flow into them only for a short time. The bilabial closure is then released with plosion. The tongue assumes its position for the next vowel throughout.

(ii) Acoustic Features.

On Mingogram 8. [sɪmbá], there is, as expected, no mouth, nasal or larynx record for the fricative [s], whereas the vowel [i] has the highest amplitude on its mouth and larynx tracing of all the segments in the utterance. Towards the end of this vowel, there is nasality, which shows the anticipatory lowering of the soft palate for the nasal [m-]. The latter sound has the longest duration of all the segments, including [-b], which it precedes in the naso-oral complex.

During the pronunciation of this nasal segment there is some periodic vibration of air in the mouth,

caused by pulmonic air which enters the oral cavity first before being diverted to the nasal chamber. Since the pronunciation of the nasal and the oral segments does not coincide in time, the [-b] does not share any of the nasality of [m-]. The final part of the vowel [a] has some very low amplitude nasal vibrations, probably the result of the anticipatory lowering of the soft palate to its position during quiet breathing.

The nasal segment [m-], as we have seen in the section on nasals, has a vowel-type formant structure. On this Mingogram this segment carries pitch at 190Hz, whereas [-b] does not have much of a pitch tracing. Both the linear and the logarithmic intensity tracings are higher and level during the nasal, whereas they show a steep fall during the plosive, immediately followed by an equally steep rise when the bilabial closure is released.

On spectrogram 97 [sɪmbá], the terminal F2 and F3 transitions of the vowel [i] before [mb], as well as the initial F2 and F3 transitions of [a], the vowel after [mb], all point towards low loci between 80Hz and 600Hz. This is as it should be since the articulation involves a bilabial stricture. The first formant of [m-] has striations with more energy than those on the voice-bar of [-b]. The release of the bilabial closure for this naso-oral sound is followed by a noise burst, which appears on the spectrum as a thin dark vertical line, lying between the base line and the middle of the frequency range.

(b) The breathy-voiced bilabial naso-oral phoneme /m̥b/.

(i) Pronunciation.

The nature and sequence of the articulatory strictures for this complex are the same as those of its plain-voiced counterpart just described: a bilabial closure, first with lowered soft palate for [m̥-], which is then raised to cut off nasal resonance during [-b̥]. The difference between these two, i.e. the complexes in (a) and here in (b), is phonatory: the first is plain-voiced as we have seen, but the second is breathy-voiced.³⁸

(ii) Acoustic Features.

Unlike the corresponding vowel on Mingogram 8, the vowel [i] on Mingogram 9 [sim̥ba], does not have anticipatory nasalisation before [-m̥ba]. The nasal segment [m̥-] has much longer duration than the oral plosive which it precedes. On the oscillogram, the larynx tracing, the oscillations showing voice for [m̥b] are wider apart than those of the plain-voiced complex, showing that the vocal cords open and close more slowly during the breathy-voiced complex. This means that there will be some difference in the steady-state pitch for these two sound types: [m̥] has a 158Hz pitch, compared to 190Hz for [m], for example. The oscillations are also wide apart for the final vowel [a].

38. I continue to use this term as defined by Ladefoged (1971, pp.6,8), without further elaboration.

During the mouth closure for the nasal there is low amplitude vibration in the oral cavity caused by pulmonic air entering that semi-active cavity before flowing into the active nasal cavity.

On spectrogram 98 [sím̥b̥a] both terminal F2 and F3 transitions of the vowel preceding, and F2 and F3 transitions of the vowel following [m̥p] point towards very low loci, as on spectrogram 97 above. The main differences between the plain-voiced naso-oral sound and its breathy-voiced counterpart are that the former has more energy than the latter; the release of the bilabial closure is followed by a weaker burst for [m̥p] than [mb], and both the nasal and the final vowel have more energy in the higher frequencies on the plain-voiced [mb] than on [m̥p]. This suggests that breathy voice is a feature of the syllable, a suprasegmental feature.

(c) The labio-dental breathy-voiced naso-oral phoneme /m̥y/.

(1) Pronunciation.

Whereas the preceding naso-oral complexes are made up of segments with a bilabial closure, [m̥y] is not. The first segment usually has a bilabial closure, while the second has a labio-dental narrowing. The place of articulation of [m̥y] seems to be a compromise between the two.

The inside surface of the lower lip moves upwards against the upper incisors, making a very firm contact. At the same time the upper lip touches the rim of the

lower lip. This stricture prevents lung air from escaping through the oral cavity.

The soft palate is lowered during the pronunciation of the nasal sound. When this has been done, the labio-dental stricture is eased; the air from the lungs is directed into the oral cavity, by a velic closure, issuing from the speech tract *fricatively*. The tongue retains its position for the following vowel throughout.

(ii) Acoustic Features.

On Mingogram 10. [m_yúrá], the nasal segment has a steady-state pitch of about 135Hz, while the oral fricative has one of around 105Hz. The nasal has vocal cord vibration throughout its duration. As can be seen on the oscillogram, vocal fold wave forms die away during [-y] whereas they increase in amplitude during the nasal. These wave forms are wide apart for both segments and for that part of the vowel [u] immediately preceded by [-y].

On spectrogram 99 [ndínódà m_yúrá], both F2 and F3 transitions of the vowel [a] before [m_y] point towards very low loci, at about 500 Hz and 600 Hz respectively. The corresponding formants of the vowel [u] after [m_y] do not seem to have any dramatic frequency movement since their centre frequencies are close to the above approximate loci.

The nasal segment has vocal cord vibrations throughout its duration, though its striations are wide apart. After this nasal but before the steady-state of the fricative, there is a voiceless interval. The voicing

of [-y] is shown by wide-apart striations, which are restricted to the lower frequency region, not exceeding 500 Hz. Above this frequency the friction of this segment is random and low in intensity compared to other fricatives (see [s] on spectrograms 97 and 98).

5.3.12.2. The alveolar naso-oral phonemes.

(a) The plain-voiced alveolar naso-oral phoneme /nd/.

(1) Pronunciation.

As can be seen on Palatogram (xii) a and b, [ndípe] and [mándá] respectively, the oral stricture for this naso-oral sound is a closure between the sides of the tongue on the one hand, and the left and right alveolar zones on the other. There is not much difference between these two Palatograms in the width of this side contact, although a high vowel is involved in the first, and a low one in the second, example.

There is a difference in the area covered by the central closure, i.e. between the alveolar ridge and the tip and blade of the tongue. Both start at about the first molar line; but the contact in example a comes very close to the lateral incisor line, whereas that in b stops in the first quarter of the canine tooth. The front of the tongue therefore interferes with the pronunciation of [nd] in [ndípe].

The sequence of other strictures is the same as in the other naso-oral segments in the preceding pages, i.e. a lowered soft palate for the nasal, raised for the oral plosive; the pressure built behind the oral closure being then released explosively for [-d].

(ii) Acoustic Features.

On Mingogram 11. [ndíró], the nasal segment is nearly three times the oral plosive in duration. The oscillogram shows that there is vocal cord vibration throughout the duration of [nd], although the wave forms have their lowest amplitude during the very last portion of the plosive element.

There are some very low amplitude vibrations in the mouth during the pronunciation of the nasal, caused by air from the lungs vibrating in sympathy with the air in the nasal chamber.

During the steady-state of the vocal cord vibration for the two segments, the nasal has a maximum pitch of 150Hz, compared to the 120Hz of the plosive. Intensity, both on the linear and logarithmic scales, rises progressively during the nasal segment, whereas it plunges downwards during the plosive, and shoots upwards immediately the alveolar contact is lost.

Spectrogram 100 [ndíró] shows that the description of this sound complex as alveolar is well-founded: the initial F2 and F3 transitions of the vowel [i] after [nd] point towards loci at about 1.7KHz and 2.5KHz respectively, as for all alveolar sounds. The nasal itself has a clear three-formant structure, F2 having the least energy.

The release of the alveolar closure for the plosive is followed by a noise burst, appearing on the spectrum as a thin vertical concentration of energy in the frequency range 500 Hz and above. There is far more energy above 2 KHz, however. This is a feature observed on spectrograms of utterances containing the breathy-voiced

plosive [ɖ]: the plain-voiced alveolar explosive is an allophone of [ɖ] occurring only with the nasal in [nd].

(b) The breathy-voiced alveolar naso-oral phoneme /nd/.

(i) Pronunciation.

The articulatory strictures for this sound are identical to those of the plain-voiced one above. The difference between them has to do with their laryngeal, i.e. phonatory stricture, [nd] being breathy-voiced.

(ii) Acoustic Features.

The most outstanding feature on Mingogram 12, [ndíró], is that the nasal segment and the oral plosive have steady-state pitches of 140Hz and 130Hz respectively, compared to 150Hz and 120Hz for their plain-voiced counterparts. The nasal segment on this Mingogram has about three times the duration of the plosive.

During the alveolar closure for the nasal, pulmonic air flows into the oral cavity causing some low amplitude wave forms - as can be seen on the mouth tracing. Also during the pronunciation of the nasal segment intensity rises steadily both on the linear and logarithmic scale. But it falls (~~less~~ rapidly on the linear than on the logarithmic scale) during the hold stage of the plosive, and rises immediately the alveolar closure is lost.

Spectrogram 101. [ndíró], also displays a clear three-formant vowel-type structure for the nasal segment, the second formant having the least amount of energy. The striations on the voice-bar of the plosive

[-ḍ] have much weaker energy than those of the plain-voiced sound. The release of the mouth closure for [ḍ] is also followed by a burst concentrated in the same frequency range as that following [-d] in [nd]. But this burst (found on all naso-oral complexes but having no apparent linguistic significance) has less energy on the breathy-voiced sound than on the plain-voiced one.

The second and third formants of the vowel [i] on this spectrogram have initial transitions pointing towards much the same high loci as after [nd] above.

(c) The "affricative" breathy-voiced alveolar naso-oral phoneme /ṇʒ/.

(i) Pronunciation.

The articulatory stricture for this sound complex is a closure between the sides of the tongue and the left and right alveolar zones, beginning at the extreme back of the palate and reaching as far forward in the mouth as the canine line - see both Palatograms (xv) a and b, [ṇʒimbe] and [wàṇʒà] respectively. Central area contact, between the tip and blade of the tongue and the alveolar ridge, extends farther back in the mouth on example a than on b, covering the whole of left and right central zones 4, touching the first molar line. On example b on the other hand, this closure is much narrower. On the whole, the width of the side closures is greater on example a than on b. These factors are accounted for by the nature of the vowels adjacent to [ṇʒ] in the two examples. In example b we have less interference from the vowel than in a.

The fact that both segments of this complex have an alveolar place of articulation means that [z̥], the dental fricative of Palatograms (ii) a and b, has been assimilated to [-n̥-], the alveolar nasal. There is, initially, an alveolar closure for the nasal, as well as a lowering of the soft palate. Pulmonic air thus flows out through the nose. But for the fricative, the alveolar stricture becomes one of narrow approximation, so that lung air, now prevented from flowing into the nasal chamber by a velic closure, issues fricatively by the mid line of the oral cavity - the closure along the sides being maintained.

Throughout, the vocal cords are positioned for producing breathy-voice.

(ii) Acoustic Features.

On mingogram 13. [n̥z̥ùngú], "pea-nuts", the two segments of the naso-oral segment seem to be about equal in duration. At the same time as lung air is vibrating in the nasal chamber, that in the oral cavity also vibrates, as can be seen by the low-amplitude wave form.

When the vocal fold vibrations have attained a steady-state for both segments, the nasal sound has a pitch of 140Hz, compared to 120Hz for the fricative. The oscillogram also shows that although the wave form for the fricative has higher amplitude than that for the nasal, its components are farther apart, hence the difference in their steady-state pitch. As can be seen on Spectrogram 102 [kè'tá n̥z̥ù'mà], the nasal segment has a formant structure and a voice-bar. The latter has striations that are closer

together than those of the voice-bar of the fricative segment. This fricative has an energy concentration in the very high frequency region above 3 KHz, as when it is in isolation, i.e. without the nasal preceding.

The second and third formants of the vowels surrounding [nɛ] have transitions pointing towards high loci at about 1KHz and 2KHz, which is higher than the loci for [z] without the alveolar nasal preceding it.

(d) The lip-rounded / 'labialised' alveolar naso-oral phoneme /nɛ/.

(i) Pronunciation.

The articulatory stricture for this sound complex is a firm closure between the sides of the tongue and the left and right alveolar zones, going as far back as beyond the fourth molar line - see Palatograms (xvi) a and b. On both Palatograms central area contact - between the alveolar ridge and the tip and blade of the tongue - is narrow, covering about half of zones 3 and 4 on the right, and the whole of zone 4 and part of zone 3 on the left central zone. There is again a marked difference in the width of side contact between the example a [nɛ^hmbò] and example b [nɛ^ha] due to the front vowel following [nɛ] in a. Here the contact is much wider than in b.

The nasal is pronounced with the soft palate lowered, allowing lung air to flow into the nasal chamber. There is a velic closure during the pronunciation of the alveolar fricative, so that lung air flows into the oral cavity. The central closure in the mouth is reduced to a narrowing, so that air flows out of the mouth with friction.

Throughout the pronunciation of this complex sound two other strictures remain constant: the bilabial, which is articulatory; and the laryngal, which is phonatory. The two lips are rounded closely and protruded, even during the pronunciation of the nasal. The vocal cords are positioned for producing breathy-voice.

(11) Acoustic Features. The nasal segment on Mingogram 14. [n₂imbò] has nearly twice the duration of the fricative. The former segment also has an oscillogram with a wave form of much higher amplitude, throughout its duration. On the other hand, the wave form of the fricative is confined to the first half of its duration, with a progressively declining amplitude. When vocal cord vibration has attained a steady-state for the nasal, that sound has a maximum pitch of 140Hz, compared to 116Hz for the fricative at a corresponding point - which is about half-way through the duration of each segment.

On spectrogram 103 [ndàná₂á], the terminal transitions of the formants of the vowel [a], preceding [n₂], and the initial transitions of the "same" vowel after this sound complex, point towards high loci. The second and third formants point towards loci at about 1.30KHz and 2.25KHz respectively.

The nasal segment has a clear formant structure, with F2 having the lowest energy. The fricative has very weak striations on the voice-bar, mainly in its latter part, next to the final vowel. The noise component

of this segment has a lower frequency limit at 2.5KHz, which is the same as when the segment is surrounded by vowels, without the nasal.

(e) The phoneme /nɔ̃/.
nasalized alveolar lateral naso-oral
phoneme /nɔ̃/.

(1) Pronunciation.

Palatograms (xvii) a [nɔ̃́] and b [nɔ̃́ŋga] show that there is a firm closure along the sides of the mouth between the left and right alveolar zones and the sides of the tongue. In example a, where the naso-oral sound is followed by a back mid vowel, the side contact is much wider than in example b, where a low vowel is adjacent to this sound. Central area closure seems to reach slightly farther forward in the mouth on example b than on a, although in both cases the stricture is alveolar.

During the pronunciation of the nasal air issues through the nasal chamber, with the soft palate in a lowered position. When it is raised for the pronunciation of the oral fricative, all the lung air flows into the oral cavity. This air has a lateral egress through a narrowing between the tongue and part of the rear part of the palate - behind the fourth molar line - see Palatograms (vii) and (viii). On Palatogram (xvii) a the lateral narrowing seems to have been made during the pronunciation of the nasal, hence the break in the lateral closure on the right alveolar zone. The vocal cords produce breathy-voice throughout.

(11) Acoustic Features.

On Mingogram 15 [nɛ́rɛ], the nasal segment has about half the duration of the fricative. The oscillogram shows that the nasal has a wave form whose amplitude starts low, and increases progressively, ending on a level similar to that of the second half of the lateral fricative. In their steady-state, both segments have a maximum pitch of 140Hz.

Spectrogram 104 [rɛ́gǎŋbò dʒákó] has two features of some interest. First, there seems to be minimal voicing for the lateral fricative: this is also reflected on the Mingogram above where the same segment has no wave form for about half of its duration. Secondly, the initial formant transitions of the vowel [o] preceded by [nɛ́] provides further evidence of the lateral egress for air during [-ɲ] being at the back of the oral cavity: the second and third formants of that vowel have initial frequency transitions pointing towards a common locus at about 2 KHz. This, as we have seen earlier, is a defining characteristic of velar articulations. It could well be that in this particular utterance, the lateral narrowing is farther back in the mouth than elsewhere because of the back vowel following [-ɲ].

The fricative element of [-ɲ] is concentrated in the frequency region 1.30 KHz to 6.50 KHz, the same range containing the noise burst of [g] and [k] in the same utterance.

The distribution of the complex phoneme /nɛ́/ seems to be limited to those environments where it is followed by the four vowels /i, e, a, o/ in Karanga. I

have not come across an example in which this phoneme precedes the vowel /u/, although Zezuru has /nɓ̀ɪnɓ̀u/, "cattle melon".

5.3.12.3. The breathy-voiced palato-alveolar naso-oral phoneme /nɓ̀/.

(a) Pronunciation.

The Palatograms (xxiv) a and b show that there is a closure along the sides of the mouth between the left and right alveolar zones on the one hand, and the sides of the tongue on the other. This closure begins at the back of the mouth and extends as far forward as the canine line on example b [m̩ɓ̀ɪnɓ̀a]; and exceeds this line on example a [nɓ̀ɪnɓ̀a].

There are two outstanding differences between these Palatograms due to the nature of the vowel adjacent to the sound complex [nɓ̀ɪnɓ̀]. In the first place, lateral contact covers a much wider area on example a than on b. Secondly, central area closure is also wider on example a than on b: on the former example, it lies between the lateral incisor line and half-way between the first molar line and the second molar line; whereas on the latter, it lies between the canine line and the first molar line.

The first segment of this sound complex is the breathy-voiced alveolar nasal [n̩], slightly retracted to assimilate to the place of articulation of the palato-alveolar affricate [ɓ̀], the second segment. During the pronunciation of the nasal the soft palate is lowered, while the central and lateral closures in the mouth are

maintained. Lung air thus has a nasal egress. This passage is blocked for the oral affricate. Lung air is thus compressed behind the palato-alveolar closure, and between the lateral closures. Lateral contact is retained while the central closure is eased slowly, so that the air in the speech tract escapes fricatively and centrally through the mouth.

Throughout, the vocal cords produce breathy-voice, and there is slight lip-rounding and protrusion.

(a) Acoustic Features.

Throughout the pronunciation of this sound complex the vocal cords vibrate producing breathy-voice, as can be seen by the wave form on Mingogram 16 [ndʒivá]. The larynx tracing shows that the components of that wave form are wider apart for [ndʒ] than for all following segments in the same utterance.

The steady-state pitch of the nasal is about 145Hz, compared to 125Hz for the affricate. Vocal cord vibration attains a higher amplitude during the nasal than during the affricate. The nasal also has longer duration than the affricate.

The nasal segment has three low-intensity formants on Spectrogram 105 [ndípe mǎndʒúg^wé]. The affricate has some striations on the voice-bar, but these are more apparent towards the end of this sound, next to the following vowel, than in the middle. The noise component of the affricate has a lower frequency limit of about 2KHz. And the second formants of the vowels surrounding [ndʒ] have transitions pointing towards a high locus at about 2KHz.

5.3.12.4. The velar naso-oral phonemes.

(a) The plain-voiced velar naso-oral phoneme /ŋg/.

(1) Pronunciation.

Both segments of this naso-oral phoneme have their articulatory stricture at the back of the mouth. The back of the tongue makes a firm closure with the velum. For the nasal segment the air passage through the nasal cavity is open since there is no velic closure. For the plosive, however, all the pulmonic air is channelled into the oral cavity behind the velar closure, the nasal cavity being cut off by a velic closure.

This air is then compressed until its pressure is higher than that of atmospheric air. Then it is released with plosion. The lips assume their position for the following vowel and the vocal cords vibrate periodically throughout.

(ii) Acoustic Features.

The nasal sound has about four times the duration of the plosive. But both segments have continuous voicing, the wave form of [g] having a higher amplitude than that of the nasal. These segments, as can be seen on Mingogram 17 [ŋgùrùwè], have each got a steady-state pitch of 140Hz.

The "same" utterance was recorded and spectrogram 106 obtained. On that spectrogram the nasal has three low intensity formants, and its voice-bar is continuous with that of the plosive. The vowel following this naso-oral complex does not have much energy on its F2 and F3.

320

(b) The breathy-voiced velar naso-oral phoneme /ŋg/.

(i) Pronunciation.

The velar strictures described for the plain-voiced naso-oral phoneme in the preceding section are identical to those required for this breathy-voiced counterpart. The only difference between these two phonemes, as their labels suggest, is phonatory: the vocal cords producing plain-voice for [ŋg] and breathy-voice for [ŋ̤g].

(ii) Acoustic Features.

On Mingogram 18 [ŋ̤gubáni], the nasal segment has twice the duration of the plosive. Although the oscillogram has a wave form throughout the nasal, there is a break in this wave form close to the end of the plosive. For both segments, the components of the wave form are wide apart, corresponding to a steady-state pitch of 130Hz and about 116 Hz for the nasal and the plosive respectively.

The utterance [ŋ̤gubáni] was also used to obtain the spectrogram 107. The striations on the formants of the nasal are wide apart, as are the few that appear on the voice-bar of the plosive. On this spectrogram the vowel [u] preceded by [ŋ̤g] has even less energy above 500Hz than on spectrogram 106. This is almost always the case with vowels preceded by breathy-voiced consonants.

5.3.12.5. The labial-velar breathy-voiced naso-oral phoneme /mbʁ/.

(a) Pronunciation.

The pronunciation of this naso-oral sound is more involved than that of those just discussed above. Whereas all the other examples are made up of segments which are homorganic or semi-homorganic, this complex, as can be seen from the symbolization, has two bilabial segments followed by a velar. Auditorily and Kinaesthetically, the elements of this complex come in a sequence although, like the others above, they function as one phonological unit.

The first two segments have the same place and manner of pronunciation as those of the naso-oral complex [mb], above, i.e. a bilabial closure which is maintained both while the soft palate is down for [m] and up for [b]. During the latter segment, pulmonic air, which previously had a nasal egress, is compressed behind the bilabial closure, with the tongue probably anticipating its position for the velar fricative [ʁ]. This is confirmed by the fact that the release of the bilabial closure for [b] sounds weaker than elsewhere. It seems the pressure build-up behind the oral closure is reduced by the narrowing between the back of the tongue and the soft palate. Instead of the explosive rush of compressed air which would follow the release of [b] in a prevocalic position, some of the air is retained for velar friction. All the segments in this complex have breathy-voice.

(b) Acoustic Features.

On Mingogram 19. [zìmbyéndé], "a cowardly person", the nasal segment [m̥] has the longest duration of the elements of [mbɣ]. That segment also has a steady-state pitch of 138 Hz, compared to 125Hz each for [b̥] and [ɣ̥]. Throughout the duration of the whole naso-oral complex, the wave form on the oscillogram has components which are wide apart.

Spectrogram 108 was obtained for the utterance [mbɣírembɣíre], "dry meal". This utterance shows the features of the second realization of [mbɣ] preceded by the vowel [e] and followed by the vowel [i]. Whereas the first vowel has a second formant with a falling terminal transition before [mb̥-], the labials, the second vowel has a common, high locus for the initial transitions of F2 and F3 at about 2.4KHz. This locus shows that the velar fricative has assimilated to the front vowel following it, becoming pre-velar.

On this spectrogram the nasal has a voice-bar continuous with that of the plosive and the fricative. Only on the first realization of [mbɣ] on this spectrogram does the weakening effect of the fricative [ɣ̥] on the plosion of [b̥] have an acoustic correlate: the release of the bilabial closure for [b̥] produces a very low-intensity noise burst, appearing on the spectrum as a thin vertical concentration of energy. The release of the second [-b̥-] on this spectrogram results in a noise burst with more energy, but spread evenly over the whole frequency range.

This phoneme does not seem to occur before the back rounded vowels /o, u/. I cannot find any Karanga

examples, and Hannan (1974) does not record any lexical items, with "mbwu" or "mbwo" at the beginning, "mbw" being the orthographic representation of the sound complex [mbɣ].

5.3.13.0. The rolled alveolar phoneme /r/.

(a) Pronunciation.

Palatograms (xxv) a and b, [rìmə] and [màràrà], respectively, show that there is a firm closure between the sides of the front and blade of the tongue on the one hand, and the left and right alveolar zones on the other. This closure along the sides of the mouth extends from the back of the palate to about half-way between the first molar line and the canine line.

There is a big difference in the width of the side contact between these two Palatograms. In example a, the area of contact is about twice as wide as that in example b. And, as can be seen on the diagram below, the soft palate remains up, cutting off nasal resonance. The lips retain their position for the following vowel.

Periodically vibrating lung air thus flows into the mouth-pharynx cavity. The tip of the tongue, which is held close to, but without touching, the alveolar ridge, is then made to flap several times against that ridge. The side contact is maintained throughout. As can be seen on the Palatograms, there is no difference in the area of contact covered by the tip of the tongue on examples a and b: it is the narrowest central area contact of all the examples described in this chapter.

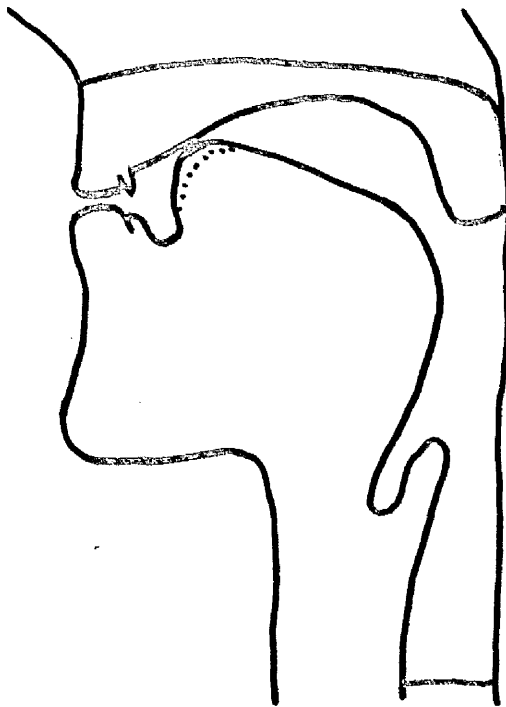


Diagram 28. /r/

(b) Acoustic Features.

Many examples used to obtain spectrographic data for this thesis contain the sound [r]. For example on spectrograms 65 [sárá múmà súrè], 71. [zàáři máʃurá], 73. [tàxórà ɲási] and 75. [ndàrúm^wá páʃùrò], the alveolar rolled sound is adjacent to all the vowel sounds of the language.

The first realization of [r] on spectrogram 65 shows that the vowel [a], both before and after the sound, has F2 and F3 transitions touching fairly high loci, at about 1.25KHz and 2 KHz. The sound [r] itself is represented by a lighter concentration of energy on the spectrum with a very short duration, about 3 centi-seconds.

Spectrogram 71 is also interesting in that it shows that the rolled [r] shares an important feature with the English r-sounds synthesised by O'Connor et al. (1957): on both realizations of [r], the F2 and F3 of the surrounding vowels having transitions which touch their loci. This, as we have seen in chapter 3, is a feature shared by the semi-vowels [w, j] and the labio-dental frictionless continuant [ɸ].

Mingograms 1. [ḍòróḅà] and 17 [ŋgùrùvè] both have a larynx wave form showing that there is vocal fold vibration throughout the duration of [r]. There are three falls in intensity (on both the logarithmic and the linear scale) perhaps corresponding to the number of closures, i.e. taps of the tip of the tongue against the alveolar ridge - see Mingogram 17. On Mingogram 1, however, there is only one fall. There would seem to be free variation in my speech between the rolled [r] and the single [ɾ].

5.3.13.1. The Labialised phonemes of Karanga.

Introduction.

In an earlier chapter (Chapter I, p181) I made reference to the so-called Cw clusters of Shona. According to Doke (1931, pp.109-124)³⁹ and Fortune (1968, 1955)⁴⁰ the phoneme /w/ combines with nearly all the other consonantal phonemes of the language and is realised by

39. Doke, C.M. (1931) op.cit.

40. Fortune, G. (1968) op.cit.

_____ (1955) op.cit.

various allophones, ranging from [x̥], to [k̥], depending on the phoneme which precedes it. I have argued that since in Karanga /x̥, ɣ̥, k̥, ɡ̥/ are phonemes in their own right, the Doke-Fortune solution may not be the best one for the dialect. The occurrence of these phonemes in such combinations as /px̥, b̥ɣ̥/, I would like to suggest, has nothing to do with /w/. Indeed, these two examples have been described in the section on affricates where no features of [w] were observed.

It has been found that (using spectrographic Electro-Aerometer and Kinaesthetic evidence) these segments have separate, sequential articulations, and are therefore not velarizations in the usual sense of the term.⁴¹ In the following pages I shall use the phonetician's primary (or primitive) tools to argue that "Cw clusters" are in fact C^w phonemes (where Cw = a sequence; C^w = a modification of C, a consonantal phoneme.) C^w phonemes are labializations of a special type: they are not environmentally conditioned. They are labialised phonemes in their own right, like the 'labialised' alveolar phonemes /ɬ̥, ʒ̥/ (pp 252-255). They are called "labialised" because the lip-rounding (a stricture of wide approximation) is simultaneous with another stricture (a closure or a stricture of narrow approximation).

41. "Phonetic labels ... are part of established phonetic vocabulary, and as such have exact and agreed definitions subscribed to by all trained phoneticians (in the ideal situation at least)." Laver, J. (1974) "Labels for voices" Journal of the International Phonetic Association, pp.62-75.

Karanga has a number of these phonemes, which are predominantly velar: /k^w, g^w, ŋ^w, ŋ̥^w, ŋg^w, x^w, ʃ^w, m^w/, where "w" is a diacritic for lip-rounding.

5.3.13.2. The labialised bilabial nasal /m^w/.

(a) Pronunciation. In addition to the lip closure as for the plain bilabial nasal [m], there is also close lip-rounding and a certain amount of lip-protrusion. The soft palate is lowered, channelling periodically vibrating lung air into the nasal chamber. The sound so produced has a peculiar auditory effect because of the lip modification.

(b) Distribution.

This phoneme occurs in three types of environment:

(i) as the class prefix of vowel-commencing stems of noun classes 1 and 3:

Class 1: /m^w-àná/, "child"; /m^w-engà/, "bride"
/m^w-enì/, "a guest".

Class 3: /m^w-ená/, "burrow", /m^w-edzí/, "moon" or "month"
/m^w-ezá/, "rat's path"; /m^w-andò/, "wind".

(ii) as the passiviser with transitive verbs:

/-rú má/, "bite" /-rú m^w á/, "be bitten"
/-mém á/, "survey" /-mém m^w á/, "be surveyed"
/-tém á/, "hit with object" /-tém m^w á/, "be hit"
/-tšém á/, "mourn" /-tšém m^w á/, "be mourned".

(iii) as non-derivative syllable onsets in nouns, verbs etc.:

/-tsám^w á/, "be angry"; /-rám^w á/, "refuse food in protest"
/vám^w é/, "others"(cl.2); /-zím^w é/, "perhaps"
/-m^w á/, "drink".

(c) Acoustic Features

On spectrograms 69. [tám^wá mášúú], 72. [ndàwánà màžùm^wí] and 75. [ndàrúm^wá páfùrò] a number of features supporting the view that [m^w] is an instance of C^w rather than a Cw may be observed. In the first place, there is only one segment, the nasal: only the features associated with [m] are visible on the spectrum. Spectrogram 72 in particular provides an ideal spectrum for comparison: in [ndàwánà], the features of [w] are the same as those described in chapter 4; these features cannot be found in that part of the spectrum corresponding to [màžùm^wí]. On this latter part of the spectrum, the labialization of [m] as [m^w] manifests itself by the abruptness of the transitions of the vowels surrounding it. All three realizations of [m^w] on spectrograms 69, 72, and 75 have this feature.

Finally, on mingogram 20, [kùtì m^wí], there seems to be no evidence to suggest there are two segments after the nasal, which would be the case if [m^wi] were pronounced and interpreted as a sequence of three segments [m-w-i].

5.3.13.3. The labialised velar phonemes.

5.3.13.4. The labialised velar plosives /k^w, g^w/.

(1) Pronunciation.

The velar stricture for these sounds is the same as for the plain plosives [k, g], i.e. a closure between the back of the tongue and the soft palate, with a velic closure. The lips, which assume their position for the following vowel in [k, g], are closely rounded for

[k^w, g^w] irrespective of which vowel is following.

Because of that lip-rounding, there is a "w"-like component to the plosion which follows the release of the velar closures for [k^w], the voiceless, and [g^w], the breathy-voiced labialised velar plosives.

(ii) Distribution.

These two phonemes occur in two types of environment:

(1) as passivisers with transitive verbs:-

/k^w/, /-ʋáká/, "build" /-ʋák^ẃ/, "be built".
/-tɕàkà/, "look for" /-tsək^ẁ/, "be hunted".
/-kóká/, "invite" /-kók^ẃ/, "be invited"
/g̣^w/, /-ʋígá/, "hide" or "bury" /-ʋíg̣^ẃ/, "be buried/
hidden.
/-gúrǎ/, "cut into two" /-gúg̣^ẃ/, "be cut..."
/-gèrà/, "cut hair" /-gèg̣^ẁ/, "have hair-cut"
/-rìgà/, "fell tall tree" /-rìg̣^ẁ/, of tall
person brought down in fight.

(2) as non-derivative consonantal onsets to syllables:

/k^w/, /k^wẽ̀zà/, "allure"; /k^wẽ̀nà/, "scratch";
 /-k^wẽ̀gùrá/, "grow old"; /mùk^wíd̥zà/, "steep part
 of road.

/g^w/, /-g^wará/, "be ill"; /mùg^wag^wà/, "road";
/tʃìg^wag^wag^wà/, "machine gun" (n.7);
/g^wenzi/, "a bush".

5.3.13.5. The labialised velar nasals /ŋ^w, ɲ^w/.

(a) Pronunciation.

The velar strictures for these two sounds are the same as those for the plain nasals [ŋ, ɲ]: i.e. a closure between the back of the tongue and the soft palate with a velic opening. Air from the lungs thus flows into the nasal chamber. The lips are closely rounded, giving "w"-sound simultaneous with the nasality while the velar closure is maintained for both the plain-voiced [ŋ^w] and the breathy-voiced [ɲ^w].

(b) Distribution.

These two phonemes occur in two types of environment:

(i) as passivers with transitive verbs:

/ŋ^w/, /-ŋéŋá/, "nibble" /ŋéŋ^wá/, "be nibbled"

/ɲ^w/, /-ɲàŋà/, "hurry for.." /ɲàŋ^wá/, "be hurried for.."

(ii) as non-derivative syllable onsets:

/ŋ^w/, /ŋ^wadí/, "letter" < Nguni; /ŋ^wará/, hill with
rich soil among rocks (n.5).

/-ŋ^wáúdzá/, "annoy by making loud noise"

/ŋ^wíró:/, ideophone of being silent.

/n^w/, /dɔŋ^wé/, "drop of rain" or "water".

/raŋ^wani/, "tall thin person" < English "long one"

/pɔŋ^weni/, name of a hill where ancestors grazed
cattle.

(c) Acoustic Features

On Mingogram 21. [wáŋ^wiwa], "it (beer) has been

drunk" the beginning and the end of the nasal segment are clear. There is a nasal wave form which coincides with an oral one: though the latter is of lower amplitude than the nasal wave form, its amplitude is higher than that of most oral wave forms which have been observed on the mingograms of nasal sounds. This could be due to the fact that the oral cavity has different resonance characteristics during the pronunciation of ordinary nasals than during the labialised ones. The vowel [i] has assimilatory nasalization.

On this mingogram [ŋ^w] has a steady-state pitch of 200 Hz, compared to 130Hz for its breathy-voiced counterpart [ŋ^w], on mingogram 22. [dɔŋ^we], "drop of rain" or "water". The nasal wave form for this latter sound has a lower amplitude than that of its plain-voiced counterpart. The oral cavity wave form also has lower amplitude and its components are wider apart than those of the plain-voiced sound. The final vowel [e] has some assimilatory nasalization.

5.3.13.6. The labialised velar naso-oral phoneme /ŋg^w/.

(a) Pronunciation.

The articulatory stricture which distinguishes this sound from the plain naso-oral [ŋg] is the secondary feature of labialization: the lips are closely rounded irrespective of which vowel is following.

Other strictures are the same for the two sound types: a velar closure accompanied by a velic opening for [ŋ-]; followed by a velic closure and the compression of

lung air behind the velar closure for [-g]. The plosion which follows the release of this closure has a labial quality because of the lip-rounding.

(b) Distribution.

This phoneme, which has no breathy-voiced counterpart, occurs in two types of environment:

(i) as the passiviser with transitive verbs:-

/-ràngà/, "discipline"(v.t.) /-ràng^wà/, "be disciplined"

/-kángá/, "fry" Ákáng^{wá}/, "be fried".

/-tʃíngá/, "deliver a baby" /tʃíng^{wá}/, "be helped to
deliver baby"

/-ʋèrèngà/, "count" or "read" /-ʋèrèng^wà/, "be read" or
"counted".

(ii) as non-derivative syllable onsets:

/tʃɪŋg^{wà}/, "bread"; /màŋg^{wà}nà/, "tomorrow";

/kùŋg^warà/, "to be clever" or "wise";

/ŋg^wenà/, "crocodile".

5.3.13.7. The labialised velar fricatives /x^w, ɣ^w/.

(a) Pronunciation.

There is a stricture of narrow approximation between the back of the tongue and the soft palate, as for [x, ʁ̥]. With the nasal cavity cut off by a velic closure, pulmonic air has two modes of flow into the pharynx-mouth cavity: aperiodic for [x^w], and periodic for [ʁ̥^w], the breathy-voiced sound. For both, the lips are closely rounded.

Whereas friction seems to be restricted to the velar region for the non-labialised $[x, \text{ɣ}]$, $[x^w, \text{ɣ}^w]$ have cavity friction, with the velar narrowing acting as the source of the noise.

(b) Distribution.

These two phonemes also occur in two types of environment:

(i) as the passiviser with transitive verbs:

$/x^w/$: $/-\text{ɓ}á\text{t}á/$, "touch" or "catch" $/-\text{ɓ}á\text{x}^w\text{á}/$, "be caught"

$/\text{ɣ}^w/$: $/-\text{ró}\text{v}á/$, "beat" $/-\text{ró}\text{ɣ}^w\text{á}/$, "be beaten".

$/-\text{rà}\text{v}à/$, "read" $/-\text{rà}\text{ɣ}^w\text{á}/$, "be read".

(ii) as non-derivative syllable onsets:

$/x^w/$: $/x^w\text{á}\text{x}^w\text{á}/$, "fat" (n.5); $/mù\text{x}^w\text{á}\text{b}á/$, "big piece of meat";

$/x^w\text{é}\text{r}è/$, "young, immature people" (n.9/10).

$/-\text{x}^w\text{í}\text{r}á/$, "pound (grain) in or for...";

$/-\text{x}^w\text{í}\text{g}á/$, help someone lift burden onto shoulder.

$/\text{ɣ}^w/$: $/\text{ɣ}^w\text{é}/$, "rock"(n.5); $/-\text{ɣ}^w\text{í}\text{p}à/$, "to glitter".

$/\text{ɣ}^w\text{é}\text{nd}\text{é}/$, "depth of river"; $/\text{ɣ}^w\text{á}\text{n}á\text{n}á/$, "little children".

$/\text{ɣ}^w\text{í}\text{r}ò/$, "grave"(n.14); $/-\text{ɣ}^w\text{é}\text{r}é\text{k}é\text{t}à/$, "talk" (v.i)

331

PART II

A STUDY OF SOME
SEGMENTAL AND SUPRASEGMENTAL
PROCESSES IN KARANGA PHONOLOGY

CHAPTER 6

6.0 THE SEGMENTAL PHONOLOGY OF KARANGA

6.1 Introduction.

In the study of the phonological structure of languages, the concern is with those properties of their sound systems which speaker-hearers must internalise in order to use their language meaningfully in interpersonal communication. The linguist's attention is now no longer focused exclusively on the physical characteristics of the sounds, but also, and in a sense, mainly, on the grammatical function of those sounds.

6.1.1 Phonology and other levels of analysis.

The division of this thesis into two apparently distinct parts is a convenient descriptive strategy, and is not meant to suggest that language is actually divided and acquired in the same way. Children learning their mother tongue acquire the semantic, syntactic and phonetic aspects of their language simultaneously. Although linguists work with levels of analysis, their object of study is unitary.

There is a mutual dependence relation between phonetics and phonology. Pike (1947)¹ wrote on this matter with vigour in the face of Hockett's (1942)² view "There must be no circularity... the line of demarcation (between levels of analysis) must be sharp." Pike (1947) asks, "If

1. Pike, K.L. (1947) op.cit.

2. Hockett, C.F. (1942) op.cit.

language actually works as a unit, with grammatical configurations affecting phonetic configurations, why should we not describe the language and analyse it in that way? If forced to do so, why pretend we are avoiding it?"

Ladefoged (1971,p.4)³ seems to pass a value judgement in the way he describes the relationship between phonetics and phonology:

"Accounting for systematic phonetic contrasts is in itself a trivial and uninteresting task that can be done in many ways. It becomes interesting only when we try to constrain our account so that it fits in with the division of sounds into the natural classes required in phonological rules."

Other writers, for example Allen and Van Buren (1971, pp.69f)⁴ place equal importance on phonetic analysis and other levels of grammar:

"Generative phonology⁵ is not an autonomous level of linguistic description but is dependent on information from other levels of the grammar, namely, surface structure and phonetics."

6.1.2 Dependence of phonology on syntax.

The morphemes of a language may be realised differently according to whether they appear in isolation

3. Ladefoged, P. (1971) op.cit.

4. Allen and Van Buren (eds.) 1971, Chomsky: Selected Readings, O.U.P.

5. This phonological study is conducted within the framework of the generative phonology paradigm as set out in The Sound Pattern of English by Chomsky and Halle. Only a few of the concepts central to this paradigm will be explained in this introduction; otherwise references will be made to relevant sections of The Sound Pattern of English, since the theory is fairly well known and my main concern here is to apply the theory to Karanga data rather than appraise some of its controversial claims.

or as part of a sentence, a nominal phrase, or some other syntactic category. An example quoted by Allen and Van Buren (1971, pp. 69-70) is the [f] of English /waɪf/, realised as [v] when combined with the plural morpheme to form a plural noun phrase /'θɪ: 'waɪvz/. Nor does /waɪf/ always change across morpheme boundaries. It does not change when functioning as part of a genitive noun phrase: /maɪ 'waɪfs 'hændpæɡ/. Similarly, the way the rules of the phonological component assign tone to the following Shona string will depend on its syntactic categorization:

/muɲu mutema/

As a declarative sentence: [muɲu mútéma] "the person is black";
 as a relative clause: [muɲu mùtéma] "a person who is black";
 and, finally, as a stabilization sentence:

[muɲú mútéma] "it's a black person".

(Part of Chapter 7 is devoted to positing some rules whereby the lexical tone of nouns of various classes is modified in stabilization constructions.)

Languages show evidence of this dependence of sound structure on syntactic information. And it is for this reason that the rules of the phonological component operate on the output of the syntactic component of the grammar.

6.1.3 Dependence of phonology on phonetics.

Transformational generative linguistic theory captures this dependence by using terminology such as "phonological representations are abstractions from phonetic data." This underlines the so-called naturalness

relation between phonetics and phonology. Phonology is concerned only with the distinctive function of sounds, whereas phonetics specifies all the physical characteristics of sounds, i.e. articulatory, auditory and acoustic features, including those features which are conditioned by the phonetic environment.

In order to show that phonological processes have a phonetic or natural motivation, we need to know a lot about the physical properties of the sounds involved. But if we meet problems in the specification of phonological processes, we turn to the phonetic theory. If the latter is found lacking, then research is conducted and reclassifications of sounds made (if need be) so that the theory becomes sophisticated enough to answer such problems. In other words,

"Phonological rules must be expressed in terms of parameters established by phonetic theory but problems of description at the phonological level may lead in their turn to a modification of phonetic theory"
(Allen and Van Buren, 1971, p.75).

6.1.4 Levels of Sound Representation.

To characterise the relationship between the systematic phonemes of a language and its inventory of phonetic segments, two levels of sound representation are postulated, the phonological representation between slashes, /..../, and the phonetic representation shown between square brackets, [....]. Since the phonological representation contains only the distinctive sound units and not the predictable redundant phonetic information, it is regarded by

generative phonologists as the level that approximates to the mental representations which speakers have of the sounds in their language.

Phonetic representations, i.e. the pronunciation of sentences, are derived from phonological representations by rules of the phonological component. Such rules are valued more if they entail linguistically significant generalizations. The postulation of one underlying form at the systematic phonemic level from which surface alternants are derived is intended to account for the intuitive knowledge which native speakers have of general or systematic relationships.

6.1.5 Linguistically significant generalizations.

The question "what is a linguistically significant generalization within transformational generative grammar?" breaks down into two major questions:

- (a) What is a generalization within this theory?
- (b) What conditions must a generalization meet to be "linguistically significant"?

In empirical science there are various types of generalizations which differ in function, form and content. Kaplan (1964 ,p.105)⁶ draws a distinction between simple generalizations on the one hand, and theoretical generalizations on the other. A simple generalization is defined by Kaplan as

"the product of a simple induction from some to all of an appropriately specified kind. A number of instances are known, and

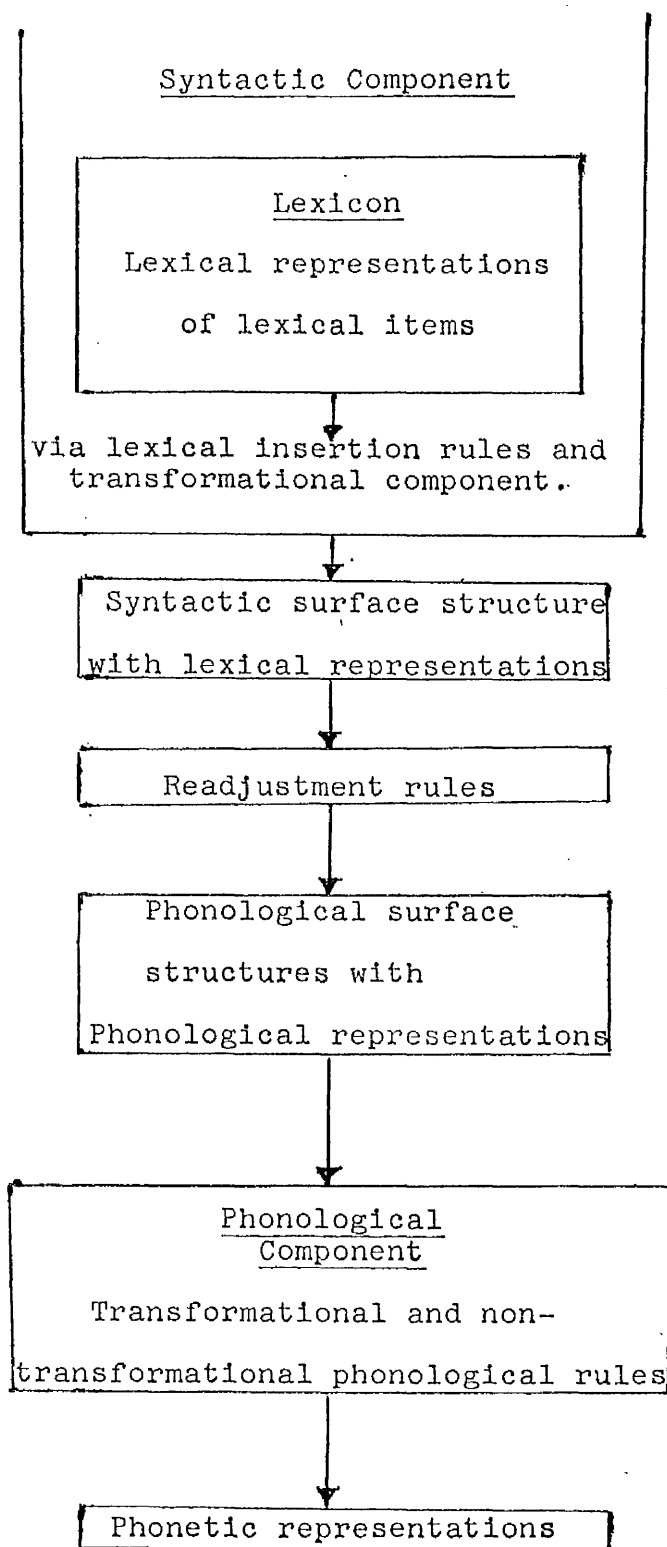
6. Kaplan, A. (1964) The Conduct of Inquiry: Methodology for Behavioural Science, San Francisco, Chandler Publishing Co.

we generalise from them to all instances of what we are prepared to call 'the same kind'."

A theoretical generalization, on the other hand, is not the product of a simple, inductive abstraction from data. As suggested by the epithet 'theoretical', such a generalization is conducted in terms of theoretical concepts that cannot be regarded as mere abstractions from the data. The distinguishing features of the generalizations encountered in transformational generative grammar are as follows (all page references are to The Sound Pattern of English): these generalizations

- (i) are "general statements" that "express regularities" (pp.330,296);
- (ii) are "compatible" or "consistent" with, the observed data (pp.330,331);
- (iii) go beyond the data in scope in so far as they deal with "potential" data;
- (iv) go beyond the data in depth in so far as they express "the facts that underlie the data" (p.330);
- (v) are expressed in terms of formal devices, i.e. in terms of theoretical concepts, and notational conventions (p.330).

6.1.6 The structure of a transformational generative grammar. (Chomsky, 1965).⁷



[.....]

7. Chomsky, N. (1965) op.cit.

6.1.7 The function of the phonological component.

The basic function of the phonological component is to relate the syntactic surface structure of each sentence to its phonetic representation or pronunciation. In the case of free variation more than one phonetic representation will be assigned to one surface structure. As distinct from the syntactic component of a grammar, containing rewriting rules and transformational rules and giving as its output terminal strings with structural descriptions, the phonological component plays no part in the formulation of new utterances: it merely assigns to them a phonetic shape.

"Given the surface structure of a sentence, the phonological rules of the language interact with certain universal phonetic constraints to derive all the grammatically determined facts about the production and perception of this sentence," (Chomsky and Halle, 1968, p.293).

6.1.8 A set of distinctive features for the sounds of Karanga.

Although distinctive feature theory has been incorporated in a specific way in generative phonology, it existed prior to generative phonology and is independent of it. But, as Allen and Van Buren (1971,p.76) note,

"The use of distinctive feature theory and the reinterpretation of the theory in order to permit the necessary abstractions made an important contribution to the development of generative phonology."

It is as well to note with Schane (1973,p.33)⁸

8. Schane, S.A. (1973) Generative Phonology, Prentice-Hall.

that there is nothing sacrosanct about the set of distinctive features presented by Chomsky and Halle in The Sound Pattern of English. Modifications of various types have been made in this thesis in order to arrive at a set of distinctive features which can best serve to describe the phonological processes of the language under investigation. What such features will have in common with those of Chomsky and Halle is that they should meet certain criteria. The criteria required in choosing features are summarised by Schane (1973, pp. 33-34):

- (i) The features have their foundation in phonetics. A feature may have primarily articulatory (e.g. [coronal]), acoustic (e.g. [sonorant]) or perceptual (e.g. [syllabic], [stress]) correlates: features must "hug the phonetic ground". (Since the ensuing choice of features has been preceded by a description, from an auditory, acoustic and articulatory point of view, of the sounds of Karanga, the features will have a phonetic motivation, A.P.).
- (ii) The features must be adequate for characterising important phonetic differences between languages.

(This means that the features used in this study must reflect the following facts: (a) that Karanga has a three-way phonetic distinction relating to glottal stricture (voiced, breathy-voiced and voiceless), whereas English, for example, has a two-way distinction (voiced and voiceless); (b) that the so-called naso-oral complexes function as one syllable onset in Karanga, but form the boundary of abutting syllables in English, A.P.).

(iii) The features must accommodate the principal allophones of a language. (As I observed in the section on nasals in Chapter 5, the palatal nasal /ɲ/ has, as expected, a high locus, and the second and third formant transitions of vowels adjacent to it point to separate loci. This happens when /ɲ/ is adjacent to the non-back vowels /i, e, a/. In the environment of back rounded vowels, however, the transitions of the vowel formants point to a common lower locus, suggesting that /ɲ/ has become retracted in place of articulation, although it remains, perceptually, non-velar. In such circumstances, what this third criterion demands is that the phonetic representation has to make use of a feature keeping the retracted /ɲ/ distinct from the velar nasal /ŋ/, A.P.)

(iv) Since features serve to categorise the contrasting segments (systematic phonemes) of a language, the features must accommodate all the necessary contrasts within the system.

(This condition has led to all the features being binary since a binary system allows one to state simply whether a segment has membership in a particular category. This requirement is linked to lexical (classificatory features) and phonological (phonological features) representations, A.P.).

(v) Segments which share phonetic traits often undergo the same phonological processes. A set of features must provide the appropriate natural classes for stating these phonological changes.

The importance of the concept of natural class to

to generative phonology is underlined by Chomsky and Halle (1968, p.335):

"No serious discussion of the phonology of a language has ever done without reference to classes such as vowels, stops or voiceless continuants... In view of this, if a theory of language failed to provide a mechanism for making distinctions between more or less natural classes of segments, this failure would be sufficient reason for rejecting the theory as being incapable of attaining the level of explanatory adequacy..."

6.2.0 The Distinctive Features of Chomsky and Halle.

The following features were proposed by Chomsky and Halle (1968, pp.298f) "to cover every inherent phonetic feature regardless of whether it plays a role in the phonetics of English." They divided the features into a number of classes.

(a) Major class features (p.302 in The Sound Pattern of English, passim)

- (i) [⁺sonorant]: all sonorant sounds are produced with a vocal tract configuration in which spontaneous voicing is possible, e.g. vowels and semi-vowels.
- (ii) [⁺vocalic]: vocalic sounds are produced with an oral cavity in which the most radical constriction does not exceed that found in the high vowels [i] and [u] and with vocal cords that are positioned so as to allow spontaneous voicing.
In producing non-vocalic sounds, one or both of these conditions is not satisfied.
- (iii) [⁺consonantal]: consonantal sounds are produced with a radical obstruction in the midsagittal

region of the vocal tract. Non-consonantal sounds do not have such an obstruction.

(b) Cavity features (pp.303ff).

- (i) [⁺coronal]: coronal sounds are produced with the blade of the tongue raised from the neutral position⁹; whereas non-coronal sounds are produced with the body of the tongue in its neutral position. Thus all alveolars, post-alveolars, palato-alveolars and palatals are [+coronal], whereas all other categories are [-coronal].
- (ii) [⁺anterior]: anterior sounds are produced with an obstruction that is located in front of the palato-alveolar region of the mouth. Thus all alveolars, post-alveolars, dentals and labials are [+anterior].

(c) Features relating to the body of the tongue (p.304).

- (i) [⁺high]: high sounds are produced by raising the body of the tongue above its neutral position. Non-high sounds have no such raising.
- (ii) [⁺low]: low sounds are produced by lowering the body of the tongue below its neutral position. There is no such lowering for non-low sounds.
- (iii) [⁺back]: back sounds are produced by retracting the body of the tongue from the neutral position.

(d) Secondary apertures (p.316).

- (i) [⁺nasal]: nasal sounds are produced with a lowered velum which allows lung air to escape through the nose; while non-nasal sounds have a raised velum, so that air escapes through the mouth.
- (ii) [⁺lateral]: this feature is restricted to coronal consonantal sounds. Lateral sounds are produced

9. Chomsky and Halle (1968) op.cit., p.300.

by lowering the mid section of the tongue at both sides or at only one side while maintaining a central oral closure. Air is thus allowed to flow out of the mouth in the vicinity of the molar teeth.

(e) Manner of articulation features (p.317).

- (i) [⁺continuant]: in the production of a continuant sound, the primary constriction in the vowel tract is not narrower than the point where the air flow past the constriction is blocked. Chomsky and Halle (1968,p.318) point out that the trilled [r] is "... a secondary effect of narrowing the cavity without actually blocking the flow of air. Consequently there is good reason to view the trilled [r] as a continuant rather than as a stop", which I shall do here.
- (ii) [⁺delayed release]: this feature applies only to sounds produced with a closure in the vocal tract. The central oral closure is released while contact is maintained along the sides of the mouth. Turbulence is generated in the vocal tract so that the release phase of affricates is acoustically very similar to the cognate fricatives.
- (iii) [⁺tense]: "This feature relates to the manner in which the entire articulatory gesture of a sound is executed by the supraglottal musculature. Tense sounds are produced with a deliberate, accurate, maximally distinct gesture that involves considerable muscular effort; while non-tense sounds are produced rapidly and somewhat indistinctly," (p.324).

(f) Source Features (p.326).

- (i) [⁺voice]: voiced sounds are produced with the vocal folds vibrating, whereas the vocal folds remain wide apart for unvoiced sounds.
- (ii) [⁺strident]: strident sounds are marked acoustically by greater noisiness than their non-strident counterparts.

(g) Other Features (pp.309, 312).

- (i) [⁺rounded]: rounded sounds are produced with a narrowing of the lip orifice.
- (ii) [⁺Distributed]: distributed sounds are made with a relatively long constriction. Bilabial sounds are thus distinguished from labio-dentals, and tongue-blade articulations from tongue-tip ones. And presumably our problematical retracted palatal nasal, [ŋ], can be distinguished from the 'pure' velar nasal [ŋ] by this feature.

Chomsky and Halle do not have features characterising prosodic or suprasegmental phenomena: "Our investigation of these features have not progressed to a point where a discussion in print would be useful" (p.329). However, since I am investigating a tone language (Chapter 7), I have to devise features for handling tonal phenomena. Since Shona uses a two-tone system, I am going to have a feature [HIGH], written in capitals to avoid confusion with the feature relating to tongue height, namely [high]. I shall also have the feature [LONG] to describe the allophones of vowels forming the nucleus of the penultimate syllables of utterances. These two features, [⁺HIGH] and [⁺LONG], represent the first modifications of (or addition to) the

distinctive features of Chomsky and Halle in this discussion. The feature relating to vowel length is predictable and is therefore introduced by redundancy rules.

6.2.1 Problems relating to the consonantal phonemes of Karanga.

Whereas the vowels of Karanga can be adequately classified using the set of features proposed by Chomsky and Halle, the consonantal phonemes present a number of problems. The problem areas in the system are of three types:

- (i) glottal stricture.
- (ii) naso-oral articulations.
- (iii) the naso-lateral /ɲ/.

In the first place, we recognised a three-way phonetic opposition in glottal stricture for the plosives: unvoiced [p,t,k], voiced [b,d,g] and breathy-voiced [p̤,t̤,k̤]. The distinctive features of Chomsky and Halle only cater for the first two categories. For the third Karanga set of sounds we need another binary feature [⁺breath], deriving from our classificatory label "breathy-voice". This means that we shall now characterise the breathy-voiced segments as being [+breath,+voice]; the unvoiced ones as [-voice], and the implosives as [+voice,-breath]. It would be redundant to specify [p,t,k], in the classificatory matrices as [-voice,+breath] since all sounds that are [-voice] are [+breath].

This solution will also be extended to all breathy-voiced segments in the system: nasals, fricatives, affricates, and naso-oral complexes.

The naso-oral complexes present a problem of another kind: how to characterise their manner of articulation. Complexes such as [mb, nd, ŋg], where both elements are [-continuant] can be characterised by that feature. But we need a feature distinguishing them from the homorganic plosives and nasals. Another feature [± naso-oral], together with the feature [± continuant] would make the necessary distinction. But there is another class of naso-oral complexes where the second element is a fricative or affricate: [nʒ, nʒ̥, nʃ, nʃ̥].

Since the nasal is, according to the Chomsky and Halle system, [-continuant], and the fricative is [+continuant], there is here an obvious parallel with the affricates, the class of sounds which are [+delayed release]. There seems to be no reason why these naso-orals should not be regarded as [+delayed release] complexes, with the qualification that [nʃ̥] should be called a [+naso-lateral] complex, to take account of the mode of egress of its second element.

The system also contains some rather awkward "affricates" [px, bʃ̥] and the naso-oral [mbʃ̥], where the elements of the complexes are not homorganic. Our phonetic description of these sounds suggested that there is a sense in which the air held behind the bilabial closure for the stops is held back by the velar narrowing for the fricatives: the plosion is not as loud as it is in 'plain' [p, b]. To that extent, these complexes may be classified together with the affricates as being [+del. rel.]. In addition to that feature, the specifications

[+anterior,+back] will be sufficient for characterising [px, b̥, m̥] as a class apart. The two breathy-voiced complexes [b̥, m̥] would be distinguished one from the other by the feature [±nasal-oral].

Finally, to characterise the class of segments which we have labelled as the labialised phonemes, namely, [m^w, k^w, g^w, x^w, ɣ^w, ŋ^w, ŋ̥^w], we need a feature [±labialised], as distinct from [±rounded]. The latter feature characterises the following segments [o, u, ɔ̥, ʊ̥, ɔ̥̄, tɔ̥, d̥], where the lip-rounding is not, as in the series above, a secondary articulatory feature: it is integral to the production of the segments. The labialised series is in contrast to homorganic, non-labialised counterparts: whereas the rounded sounds do not have non-rounded homorganic counterparts. The two features [±labialised] and [±rounded] are being used in this special sense to characterise these two articulatory distinctions.

On the following distinctive feature matrices a combination of Chomsky and Halle's features and those proposed for some of the problematical segments of Karanga is used; and the feature [±vocalic] is replaced by [±syllabic]. I shall also use a feature [explosive] to distinguish [b̥, d̥] from [p̥, t̥]. The matrices are fully specified.

A Distinctive Feature Matrix for the Consonantal Segments of Karanga (2).

[illegible]

* Schane (1973, p.22) reports a similar distinctive feature for Twi and Nupe, two West African languages.³

6.2.2 Some phonological processes in Karanga.

Fortune (1955, pp.19-43)¹⁰ is a study of some of the phonological processes that occur in Shona (mainly in the Zezuru dialect). Although he quotes from that dialect, the processes he describes occur in the other dialects as well. On page 19 paragraph 45 Fortune sets out very clearly the form and content of the statements which he is about to make, and I quote with approval:

"In such a study (i.e. a phonological study) we are concerned with three things:
 1. the nature of the sounds that undergo change;
 2. the nature of the sounds which are the products of the changes;
 3. the conditions under which change takes place."

As a statement of the concerns of a traditional phonological study Fortune's words cannot be faulted: the only difference between the above and the transformational generative phonology formula for phonological processes is one of brevity:

$$\begin{array}{ccccc} A & \longrightarrow & B & / & X \longrightarrow Y \\ SD & & SC & & ENV \end{array}$$

This formula states that A, the structural description of a string, is realised as B, the structural change, in the environment X to the left, Y to the right, where X and Y are strings not affected by the change, and either may in fact be null. SD is the input, and SC the output of the rule.

The contribution of generative phonology in this respect is, however, more substantive than the provision of a semi-mathematical formula for old concepts.

10. Fortune, G. (1955) op.cit.

In a phonological study in generative terms SD is the explicit and unambiguous specification of the string undergoing change using distinctive features, the so-called "bundle" of simultaneous features. Similar SD's are grouped together into natural classes since they undergo similar processes. Similarly the SC has to be characterised by either a change in the feature composition or feature values of the string SD. The completeness of such statements therefore depends on the extent to which our features are based on phonetics: phonological processes are natural. Thus Postal (1968, p.56)¹¹

"... the categorization of lexical items given by phonological structure, i.e. required to represent morphemes in the dictionary, required to state morphophonemic and phonological rules, needed to state constraints on sequences of phonological elements ... is not, from the point of view of phonetic structure, an arbitrary code. Rather, this representation is closely related to the representations needed to state the phonetic properties of the various sequences which represent individual lexical items..."

In other words, the features used in classificatory matrices in the lexicon are related to those used in phonetic representations: there is a naturalness condition between phonological and phonetic representations, made explicit in phonological rules.

6.2.3 The lexical representation of Karanga words.

The lexicon of a language is a list of the lexical items - words, grammatical and other lexical formatives - of the language concerned. Thus the lexicon

11. Postal, P.M.(1968) Aspects of Phonological Theory Harper and Row, New York.

of Karanga will contain semantic, syntactic and phonological information about each item.

Since this is a phonological study, we are interested mainly in the phonological characteristics of items as represented on the classificatory lexical matrices. When the term "lexical representation" is used again in this discussion, it is to be understood as referring only to the phonological properties of a formative.¹²

The views expressed by Derek Fivaz (1970)¹³ on the lexical representation of Shona words are generally the same as mine. The phonological representation of items incorporates information about their tone pattern. The lexical tone pattern of nominal stems is unpredictable and each one has to be specified for tone - although all noun class prefixes are [-HIGH] and therefore do not need to be marked. Verb stem tones, on the other hand, are largely predictable (Fivaz 1970, pp.225-227; Fortune, 1968, Vol.II, pp.7f). The infinitive inflection is the most convenient to take as basic. This can be used as the input to the phonological component, where the tone pattern required in other inflections may be derived by rules. Generally speaking, all verb stems with initial [-HIGH] tone have all their other syllables on [-HIGH].

12. Transformational phonological rules (Chomsky and Halle, 1968, pp.15,59f) take the syntactic categorization of an item into account before being applied. So information on the syntactic characteristics of items cannot be excluded completely in a phonological study.

13. Fivaz, D. (1970) Shona Morphophonemics and Morphosyntax University of the Witwatersrand Press.

Those with an initial [+HIGH] tone, on the other hand, will have all but the fourth syllable on [+HIGH]:

$$-H_1-H_2-H_3-H_4\cdots-H_n$$
$$+H_1+H_2+H_3-H_4\cdots-H_n$$

Fivaz¹⁴ suggests that all that is needed is a simple indication of [+HIGH] or [-HIGH] for any particular verb stem in the lexicon. Redundancy rules (Chomsky and Halle, 1968, pp.171f; 380f) can then spread the tone to all the syllables of the stem in accordance with the constraints of the two stem types above. This sounds economical enough.

6.3.0 Word Derivation Processes.

In this chapter our primary concern will be to posit some rules which, it is hypothesised, the speakers of the language use in order to augment the vocabulary of their language. The focus of attention will be on the commutation of phonological features in the derivation of words of one class from those of another.

Since the influence of the class 9/10 nasal prefix on consonants at the beginning of verbal, nominal and adjectival stems is one of the most widespread processes of word formation in which phonological features are commuted, a detailed description of it will be given. Doke (1931, pp.62-71) and Fortune (1955) concentrated in their study of the phenomenon, on nominal and adjectival stems.

14. Fivaz's publication contains an extensive discussion of verbal tone in Shona in generative phonology terms (see his Appendix II pp.224-239). In this thesis we shall concentrate on nominal tonology.

This study of noun derivation will therefore focus on those nouns of classes 9/10 and 5 which are derived from verb stems. The division of Shona nouns into classes is discussed at length by Fortune (1955, pp.51-52) who defines a noun class thus,

"Taking the prefix and set of concords together as a criterion of classification, we may define a noun class as a group of nouns which do not differ in prefix and which govern the same concords."

Since the grammatical relationship holding between nouns of certain consecutive classes is one of a singular form and its plural, we shall also posit some rules to account for the derivation of the one from the other. The singular form is the one that would be represented in the lexicon, while the plural is derived from it by rules of the phonological component. For example, in the syntactic surface structure (Chomsky and Halle, 1968, pp.9-11, 371-72) the nouns of class 5 will be labelled $N_5 + \{PLURAL\}$, the grammatical formative $\{PLURAL\}$ having no phonological shape at that stage in the derivation. Thus $\begin{bmatrix} \text{ǃáǃǃǃá} \\ N_5 \end{bmatrix}$ ("knife") + $\{PLURAL\}$ will have that abstract morpheme replaced by the bundles of features abbreviated alphabetically by /mà-/ by readjustment rules, giving $\begin{bmatrix} \text{ǃáǃǃǃá} \\ N_5 \end{bmatrix} + \begin{bmatrix} \text{mà} \\ N_5 \end{bmatrix}$. Since in Shona the plural morpheme is prefixed rather than suffixed to the noun stem, the rules will prepose the plural morpheme to the stem:

$\begin{bmatrix} \text{mà} + \text{ǃáǃǃǃá} \\ N_6 \quad N_6 \end{bmatrix}$

Transformational phonological rules will then convert this phonological surface structure into its correct phonetic

[6] - commencing class 5 nouns (see 6.3.5).

Rule 3. ǂ —————→ p/màt ——— V...]_{N6}, i.e. [màpángá]
(where "V" is the first vowel of the stem,
preceded by the initial consonant of the stem).

There is also the tonal rule which says that the addition of the low tone plural/class 6 prefix to a class 5/singular stem in this particular utterance does not result in any change in the lexical tone pattern of that stem.

One must, however, add that unless the rules deriving the phonetic representations are correctly ordered, those representations will be incorrect.

In chapter 1, a problem arising from Fortune's (1955) failure to work within the framework of a well-defined phonetic theory was mentioned: namely the problem of not making use of underlying regularities to posit rules accounting for superficially disparate processes. There is the risk, when one is not making use of the concept of natural class, of grouping together items that

15. Schane, S.A. (1973) op.cit.

have little in common and making general statements about them. The two processes of "bilabialisation" and "supradentalisation" (Fortune, 1955, pp.36-37) are discussed as though they had nothing in common. The following examples of bilabialization are given:

- (1) /-ʋóná/ "see" → [mbóní], "pupil of eye"
 /-yùtá/ "blow" → [m̥yùtò] "bellows"
 /-pùká/ "animal" (Bantu) → [m̥hùká] "animal"
 /-ḡèrèkà/ "carry baby on back" → [mbèrèkò] "apron for carrying baby on back!"

(I have added slant brackets and tone marks, not found in Fortune. The derived nominals have been put in square brackets. The tone pattern of these nouns is usually the same as that of the verbs from which they are derived. A summary of the rules governing the tones of the nouns is given in Appendix I, pp.446-448).

Examples of the processes called bilabialization and supradentalization by Fortune are numerous, as the following lists will show:

- (ii) N-+/-pòndà/ "kill" → [m̥hòndì],¹⁶ "murderer"
 N-+/-péngá/ "be mad" → [m̥hèngó] "madman"
 N-+/-pópótá/ "shout at" → [m̥hópótò] "slanging match"
 N-+/-pòmbà/ "commit adultery" → [m̥hòmbɔ̃ɛ] "adulterer(ess)"
 N-+/-páridzá/ "preach" → [m̥hárídzò] "sermon"
 N-+/-pìndùrà/ "answer" (v.t) → [m̥hìndùrò] "answer" (n.9/10)

Examples involving alveolar sounds are also widespread:

16. There seems to be no regular rule accounting for the change from the final vowel of the verb to that in the final syllable of the derived noun.

- (iii) N-+/-témá/ "cut" → [nḡémó] "chisel"
 N-+/-tétémá/ "greet ceremoniously" → [nḡétémò] "praise poetry"
 N-+/-túm^wá/ "be sent" → [nḡúm^wà] "messenger"
 N-+/-tàùrà/ "speak" → [nḡàùrò] "Talks" (n.10)
 N-+/-táp^xá/ "be captured" → [nḡáp^xà] "captive(s)"

The process also affects verb stems commencing with a voiceless velar plosive; with a minor modification:

- (iv) N-+/-kóká/ "invite" → [ḡókà] "beer party for work"
 N-+/-kùjà/ "grind" → [ḡùjò] smaller of 2 grinding stones
 N-+/-kúrukúra/ "discuss" → [ḡúrukúró] "discussion(s)"
 N-+/-kùndà/ "gather refuse" → [ḡùndí] "chaff" for disposal"
 N-+/-kórómórà/ "acquire large amount" → [ḡórómórì] "professional criminal"
 N-+/-kùsùrà/ "beguile" → [ḡùsùrì] "scrounger"

For all the three places of articulation for plosives, the derivation of class 9/10 nouns seems to take place in two stages. First, the nasal prefix, N-, of classes 9/10 becomes homorganic with the plosive at the beginning of the verb stem:

Rule 4

$$[+nasal] \rightarrow \left[\begin{array}{c} \alpha \text{ ant} \\ \beta \text{ cor} \end{array} \right] / \# \left[\begin{array}{c} C \\ \alpha \text{ ant} \\ \beta \text{ cor} \\ -\text{cont} \\ -\text{voice} \end{array} \right] \dots \left. \vphantom{\left[\begin{array}{c} C \\ \alpha \text{ ant} \\ \beta \text{ cor} \\ -\text{cont} \\ -\text{voice} \end{array} \right]} \right] \text{Verb}$$

(This rule applies only to the examples marked (i), (ii), (iii) and (iv) above). This rule, which we shall call the nasal assimilation rule, gives representations at the

beginning of each of which there is a complex of the nasal and the unvoiced plosive. /mp..., nt..., ŋk/. But nasals do not combine with unvoiced consonants in Karanga.

The second rule adds breathy voice or h-quality to the nasal. The breathy-voiced nasal has been represented as N_{h} instead of N_{v} to facilitate the rule affecting the velar plosive (see below). Another way of seeing the second rule is as a process whereby the voiceless plosive becomes the breathy-voiced glottal fricative:

Rule 5.

$$\left[\begin{array}{l} \text{-cont} \\ \text{-voice} \\ \text{-del.rel.} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{+cont} \\ \text{+voice} \\ \text{+low} \\ \text{+breath} \end{array} \right] / \# \left[\begin{array}{l} \text{C} \\ \text{+nasal} \end{array} \right] \text{---} \dots \quad \text{N9/10.}$$

This rule yields representations written in square brackets for the examples (i) - (iii) above. But for the examples (iv) the representations derived are not Karanga forms:

e.g. /* ŋ_{h} oka, * ŋ_{h} ujo/.

A third rule, which we may call the velar nasal deletion rule, derives the phonetic representation required:

Rule 6.

$$\left[\begin{array}{l} \text{+nasal} \\ \text{-ant} \\ \text{-cor} \\ \text{+voice} \end{array} \right] \rightarrow \emptyset / \# \left[\begin{array}{l} \text{+cont} \\ \text{+voice} \\ \text{+breath} \\ \text{+low} \end{array} \right] \text{---} \dots \quad \text{N9/10.}$$

The reason for the deletion of the velar nasal here has, I believe, to do with the fact that the segment $[\text{ŋ}]$ or, as written here, $[\text{ŋ}_{\text{h}}]$, is rare in non-borrowed words in Shona. Karanga of the midlands region has many words with

[ŋ]-commencing syllables borrowed from the Nguni group of languages:

<u>Karanga</u>	<u>Ndebele/Zulu</u>
/-m̩aŋàrà/	/-m̩aŋàlà/ "report a crime"
/-v̩aŋà/	/-v̩aŋà/ "be in a hurry"
/ŋ̩úzu/	/ŋ̩úzu/ "a type of dance"
/-ŋ̩òrà/	/-ŋ̩òlà/ "be unclean"
/-r̩uŋà/	/-l̩uŋà/ "become disciplined".

The importance of the rules just posited lies in their utilization of the concept of natural class. The nasal assimilation rule replaces the processes called "bilabialization", "frication" and "supradentalization" by Fortune (1955, pp. 36-37) by taking into account those features shared by [p, t, k], the class of segments which are [-cont, -voice, +breath].

It will have become evident by now that the rules which I have posited above do not account for all the examples given by Fortune in (1) above. The examples which I have (deliberately) omitted are those derived from underlying verb stems which commence with the following sounds: [v̩, ɓ̩, d̩, y̩]. In terms of this theory, only those strings which have the same structural description, and undergo the same phonological processes can be grouped together and accounted for by the same rules. The rules of so-called "bilabialization", "supradentalization" and "frication" as formulated by Fortune do not account for why the voiceless plosives [p, t, k] yield respectively [m̩h̩, n̩h̩, and ɲ̩] - commencing class 9/10 nouns; why [y̩] alone yields a [m̩y̩] - commencing class 9/10 noun, and, finally,

why [ɸ, ɓ, d] yield, respectively [mb, mb, nd] - commencing class 9/10 stems.

The point I am making is this: the segments [p, ɸ, ɓ, y] have some bilabiality in common. So that is perhaps why Fortune attempted to account for them all by one "bilabialization" rule. But there are some internal differences among them which anyone writing or learning a Shona grammar must take into account:

- | | | | |
|-----|------------|-------|-----------------------|
| (1) | [ɸ] → [mb] | } ... | [+voice]
[-breath] |
| | [ɓ] → [mb] | | |
| (2) | [y] → [my] | ... | [+voice]
[+breath] |
| (3) | [p] → [mɸ] | ... | [+voice]
[+breath] |

Anyone learning the language for the first time would be corrected if they were to derive [mb]-commencing stems from [p] or [v], for example:

N-+/-pòndà/ "kill" → *[mbòndì]
 N-+/-yùtà/ "blow" → *[mbùtò]

The asterisked examples are not Shona words, although they do not violate the morpheme structure conditions of the language: they are accidental gaps.

There is another, more important reason why they are unacceptable forms. The processes (1) - (3) above fall into two natural classes: [ɸ, ɓ, d], the plain-voiced group; [p, y], the [+breath] group. Addition of the [+voice] N-prefix to the former class yields plained-voiced consonant-commencing stems; while it yields breathy-voiced ones when added to the latter class. In the former class we may

segment and derives the following representations:

(v)(a) /mʷóná/, /mʷimbà/ etc.

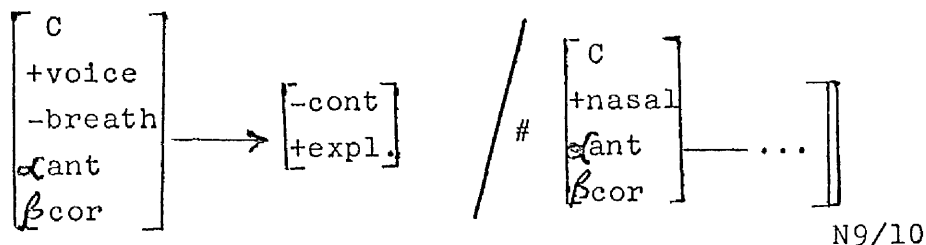
(b) /mbátʷá/, /mbèrèkà/ etc.

(vi)(a) /ndíkì/

(b) /nrímá/, /nrúmúrè/ etc.

In all cases, the nasal agrees in voicing with the segment at the beginning of the verb stem. A second assimilation rule then applies: since [ʷ, b, d, r] do not combine with other consonants in the system, they are replaced by homorganic plosives:

Rule 7.



Apart from the final vowels, which do not seem to be determined by any regular rule, and have to be accepted as given, the above rule derives the correct phonetic representations on the right hand columns in examples (v) and (vi).

The [y̥]-commencing underlying verb stem is the only one given by Fortune that yields a [m̥y̥]-commencing class 9/10 noun stem. But [y̥] is not the only segment that behaves in this way in the language. It belongs to a different class of segments which are pronounced at the lips, and at the alveolar ridge: they all yield breathy-voiced naso-oral-commencing stems:

- (vii) N-+/-yùtá/ "blow" → [m̥yùtò] "bellows"
 N-+/-byùma/ "agree" → [m̥yùmo] "permission"
 N-+/-byùvìrà/ "add yeast" → [m̥yùvìrò] "yeast"
 N-+/-byòngódzá/ "stir water" → [m̥yòng^{wé}] "muddy water"
 N-+/-byárúrá/ "tear" → [m̥yáru-m̥yáru] "bits and pieces"

But the process whereby a breathy-voiced fricative or affricate changes to a breathy-voiced homorganic naso-oral complex when the N-prefix is added to a verb of which it is the initial segment is not limited to the labials:

- (viii)(a) N-+/-zárá/ "give birth" → [nzárá-kam^{wé}] "a mammal
 with one off-spring"
 N-+/-dzòmbòrà/ "steal" → [nzòmbòrì] "pickpocket"
 N-+/-dzòpó/ ideo. of pulling back skin [nzòpó]
 "circumcision"
 N-+/-dzéngé redzá/ "cover up a mistake"
 → [nzéngéredzò] "speech to divert suspicion"
 (b) N-+/-bírá/ squander someone's money → [nbíré]
 "exploiter"

Two rules are needed to derive the phonetic representations on the right. The first rule will change the affricates into their corresponding fricatives:

$$\begin{array}{lcl} [by] & \longrightarrow & [y] \\ [dz] & \longrightarrow & [z] \end{array}$$

Then the nasal assimilation rule will apply:

$$N-+/-y.../ \longrightarrow [m̥y]; \quad N-+/-z.../ \longrightarrow [n̥z]; \quad N-+/-b/ \longrightarrow [nb]$$

Using features:

$$[+del.rel.] \longrightarrow [+cont.] / \# \text{ --- }]_{\text{verb}}$$

This rule, affecting only the affricate-commencing verb and ideophonic stems, derives the following representations:

(vii) $/-b\dot{y}\dot{u}m\grave{a}/ \longrightarrow /y\dot{u}m\grave{a}/$ - (interesting since latter $/-bv\dot{u}m\grave{a}/$ ulu/Ndebele verb from which we have borrowed $/-bv\dot{u}m\grave{a}/$)

$/-b\dot{y}\dot{u}\dot{\vartheta}\dot{i}r\grave{a}/ \longrightarrow /y\dot{u}\dot{\vartheta}\dot{i}r\grave{a}/$

(viii) $/-d\dot{z}\dot{o}mb\dot{o}r\grave{a}/ \longrightarrow /z\dot{o}mb\dot{o}r\grave{a}/$

$/-d\dot{z}\dot{e}ng\dot{e}r\acute{e}d\dot{z}\grave{a}/ \longrightarrow /z\dot{e}ng\dot{e}r\acute{e}d\dot{z}\grave{a}/$

Finally, the nasal assimilation rule applies, giving the correct phonetic representations in examples (vii) and (viii) above.

Rule 8.

$$[+nasal] \longrightarrow \left[\begin{array}{c} \alpha \text{ ant} \\ \beta \text{ cor} \\ \Sigma \text{ round} \end{array} \right] / \# \left[\begin{array}{c} \alpha \text{ ant} \\ \beta \text{ cor} \\ \Sigma \text{ round} \\ +\text{cont} \end{array} \right] \dots \quad \text{N9/10}$$

The inclusion of $[\vartheta, \beta, p, y]$ under one rule - "bilabialization" - without explaining important internal differences among them (see p.360 above), like the inclusion of $[t, d]$ under one rule - "supradentalization" - , also without explaining why

$/d/ \longrightarrow [nd]$

while

$/t/ \longrightarrow [n\grave{h}]$

is grammatical analysis at the level of observational adequacy: it does not give a principled explanation of what the processes under discussion entail. Phonological rules can be formulated using phonemes as unanalysable wholes, as Fortune did, at the risk of obscuring those

traits which such phonemes share, and do not share. So important is the concept of natural class to generative phonology. Thus, although Fortune, in his formalization of what is involved in phonological processes (1955, p.19 para 45), speaks of the nature of the sounds that undergo change, and of the nature of the sounds that result from this change, his notion of the "nature" of sounds seems to be founded on an inadequate phonetic study.

6.3.2 The Derivation of Class 5 nouns from /p,t,k/-commencing verb stems.

The significant linguistic generalizations achieved by rules deriving nouns from the infinitive inflection as the underlying representation are not limited to class 9/10 nouns. Similar generalizations characterise the rules whereby class 5 nouns are derived from certain verb stems.

Corresponding to the three places at which plosive consonants are pronounced in the language - the bilabial, alveolar and velar - three sets of class 5 nouns may be derived from verbs whose stems commence with unvoiced plosives.

- (ix) /-pákátá/ "carry under arm" → [ḡákát^wa] "big knife"
 /-pákúrá/ "serve food" → [ḡákúró] "gift to mother of bride"
 /-pámhá/ "plunder" → [ḡámhá] "bullying"
 /-pémberá/ "celebrate" → [ḡémberò] area in which dancing takes place.
 /-pàràḡḡa/ "destroy" → [ḡàràḡḡo] "destruction"
 /-péréká/ "hand back love token" → [ḡérékò] "token payment for divorced wife"

- /-pèsàná/ "disagree" → [bèsán^wà] "confusion"
- (x) /-tàm̀bù̀d̀zà/ "torment" → [dàm̀bù̀d̀zò] "vexation"
- /-táng^wá/ "be started" → [dáng^wè] "first born"
- /-tám̀bá/ "play" → [dám̀bè] "practical joke"
- /-tándád̀zà/ "entertain others" → [dándád̀zò] "pastime"
- (xi) /-kù̀pù̀rà/ "send wife back to her parents" → [gù̀pù̀rò]
- "token of divorce"
- /-kángá/ "fry" → [gángó] "frying pan"
- /-kà̀nà/ "mix" → [gà̀nò] "soil mixed for plastering"
- /-kà̀rù̀kà/ "be startled" → [gà̀rù̀kò] "fear"
- /-kà̀rù̀d̀zà/ "cause to be startled" → [gà̀rù̀d̀zò] "talent for shocking people"
- /-kú̀ngú̀rá/ "regret" (v.i.) → [gú̀ngú̀rà] "regret"

What seems to be common to all the examples above is the voicing of the plosive at the beginning of each verb stem:

Rule 9.

$$\left[\begin{array}{l} \text{-cont} \\ \text{-voice} \end{array} \right] \longrightarrow [+voice] / \# \text{---} V \dots \dots \dots \left. \vphantom{\left[\begin{array}{l} \text{-cont} \\ \text{-voice} \end{array} \right]} \right]_{N5}$$

This rule produces the following intermediate stage:

- (ix) /bákátá, bákúrá etc./
- (x) /dàm̀bù̀d̀zà, dáng^wè/
- (xi) /gù̀pù̀rà, gángá/.

But since the plain-voiced bilabial and alveolar plosives occur only in naso-oral complexes, the segments [b] and [d] at the beginning of the stems in (ix) and (x) above have to be shown to be implosive, and so another rule is required:

Rule 10.

$$\begin{bmatrix} +\text{expl.} \\ +\text{ant} \end{bmatrix} \longrightarrow [-\text{expl.}] \quad / \quad \# \quad \text{---} \text{V....} \quad \Bigg]_{N5}$$

There are no velar implosives in Karanga, so the above rule does not apply to [g]-commencing representations. Instead, a rule changing the plain-voiced velar plosive into the breathy-voiced segment operates, since [g] does not occur otherwise than in the naso-oral complex [ŋg] in Karanga:

Rule 11.

$$\begin{bmatrix} -\text{ant} \\ -\text{cor} \\ -\text{cont} \\ +\text{voice} \\ -\text{breath} \end{bmatrix} \longrightarrow [+ \text{breath}] \quad / \quad \# \quad \text{---} \text{V....} \quad \Bigg]_{N5}$$

Writers on Shona believe that there was a class 5 prefix /ri-/ which has been lost. This is corroborated by the fact that nouns of this class have a grammatical concord of the form *m/rV/*, i.e. /ri-, ro-, re-, ra-, ru-/, depending on the syntactic environment.

6.3.3 Three derivations of plural forms by (i) altering the vowel, (ii) adding a prefix and (iii) changing a consonant.

The addition of the plural morpheme to the stems of nouns of various classes in Karanga results in a number of interesting phonological changes. In this section I shall formalise the processes required to derive the plural forms of the nouns of classes 3,5 and 7. It is important

to remember that the plural morpheme manifests itself in various ways. The three noun classes to be discussed here represent three ways in which the plural morpheme is realised: by vowel alternation; by the addition of a prefix to the singular form, or by a change in the feature composition of a consonant.

Where class 9 nouns are concerned, the "derivation" of the plural, i.e. class 10, noun does not cause any phonological changes. Rather, whether the noun concerned is singular or plural is seen in the grammatical concord associated with it.

6.3.3.1 Prefixes in Primary and Secondary Function.

As Fortune (1955, pp.54-55) points out, most nouns in the language are associated with at least two prefixes - one indicating the singular and the other the plural form. For example, the stem /-kòv̀à/ may have the prefix /r̀ù-/ of class 11, or the prefix N- of class 10, the former indicating the singular while the latter indicates the plural:

/r̀ù-kòv̀à/ "river"

/N-kòv̀à/ "rivers"

These two seem to be the prefixes with which the stem /-kòv̀à/ is primarily associated. However, since this stem is also associated with other prefixes with different semantic significance, it is necessary to distinguish between the two types of prefix. Fortune quotes the following examples:

/k̀a-kòv̀à/ "small river", cl.13.

/t̀ù-kòv̀à/ "small rivers", cl.12.

The prefixes /rù-/ and /N-/ are called "prefixes in primary function", as opposed to /kà-/ and /tù-/, which Fortune calls "prefixes in secondary function"; and the relationship holding between the prefixes and the stem are called "relationship of primary function" and "relationship of secondary function", respectively.

In this discussion, the singular prefixes, and the plural prefixes derived from them have a relationship of primary function with the stem. To include the other prefixes would make this discussion inordinately long. In any case, the rules posited for the derivation of the plural forms using primary prefixes will not be fundamentally different from those which would be required to account for prefixes in secondary function.

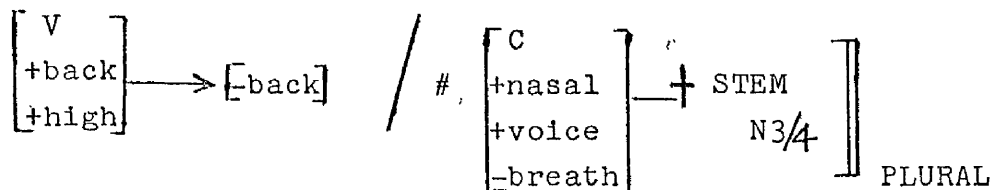
6.3.3.2 Plurals formed by vowel alternation, i.e. class 3 to 4.

Of the noun classes chosen to illustrate the process of plural form derivation, only classes 3 and 4 are related by an alternation in the vowels associated with the class prefix. Class 4 is the numerical plural of the nouns of class 3. The following are some of the words illustrating this process:

(xii)	<u>Class 3</u>	<u>Class 4</u>	<u>English</u>
	/mù-ndà/	→ [mì-ndà]	"field(s)"
	/mù-nàmàtò/	→ [mì-nàmàtò]	"prayer(s)"
	/mù-ŋòzòwá/	→ [mì-ŋòzòwá]	"name of tree"
	/mù-ŋá/	→ [mì-ŋá]	"home(s)"

The plural forms are derived by changing the vowel of the prefix from /u/ to [i]:¹⁷

Rule 12.



6.3.4.0 Plurals formed by changes in the feature composition of the consonant of the prefix, i.e. class 7 to 8.

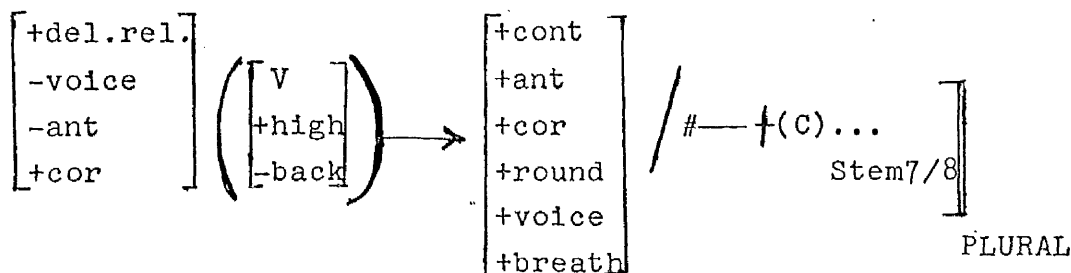
The prefixes which have a relationship of primary function with the nouns of classes 7 (singular) and 8 (plural) have the shape /tʃi-/ and /z̥i-/, respectively, before consonant-commencing stems, and /tʃ-/ and /z̥-/ before vowel-commencing stems. But this does not affect the process of plural form derivation since the latter involves only a consonantal segment change:

(xiii)	<u>Class 7</u>	<u>Class 8</u>	<u>English</u>
	/tʃi-pó/	→ [z̥i-pó]	"gift(s)"
	/tʃ-urù/	→ [z̥-urù]	"ant hill(s)"
	/tʃi-gàrò/	→ [z̥i-gàrò]	"stool(s)"
	/tʃi-timà/	→ [z̥i-timà]	"train(s)"
	/tʃ-ápupù/	→ [z̥-ápupù]	"witness(es)"
	/tʃ-otó/	→ [z̥-otó]	"fire place(s)"

The change from singular to plural is effected as follows:

17. It will have become clear by now that the phonetic shape of the plural morpheme in Karanga is different to that of other languages in that it does not seem to be conditioned by surrounding sounds. In English this morpheme may be realised as {s} in /kæt/ ~ [kæts] because it is preceded by the voiceless [t], for example. In Karanga the form of the morpheme is given.

Rule 13.



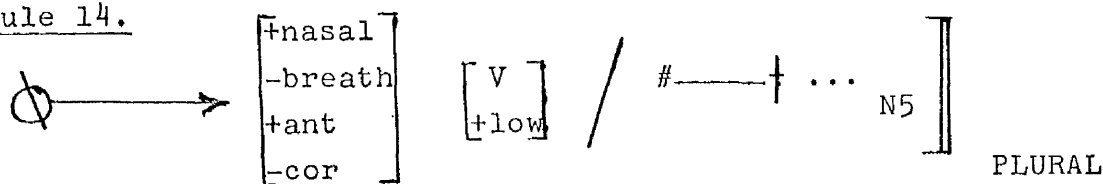
6.3.5 Plurals formed by the addition of a prefix, i.e. class 5 to class 6.

The derivation of the plural form of the nouns of class 5, i.e. the derivation of class 6 nouns, is effected in three stages: first the plural prefix is added to the underlying singular (class 5) form, then the initial consonant of that underlying form is explosivised and devoiced (if it is an implosive):

(xiv) (a) <u>Class 5</u>	<u>Class 6</u>	<u>English</u>
/bàngá/	→ [mà-pàngá]	"knife"/"knives"
/bàngò/	→ [mà-pàngò]	"pole(s)"
/bèrè/	→ [mà-pèrè]	"hyena(s)"
/bápú/	→ [mà-pápú]	"lung(s)"
(b) /dùrà/	→ [mà-tùrà]	"grain-store(s)"
/dándá/	→ [mà-tándá]	"big pole(s)"
/dèng ^w aná/	→ [mà-tèng ^w aná]	"basket(s)"
/dùndùrù/	→ [mà-tùndùrù]	name of fruit

In the first place, the plural morpheme is preposed to the underlying singular form:

Rule 14.



This rule produces the following representations:

$$\begin{array}{l} /m\grave{a}-/ + /d\acute{a}nd\acute{a}/ \longrightarrow /m\grave{a}-d\acute{a}nd\acute{a}/ \text{ etc.} \\ /m\grave{a}-/ + /b\acute{a}ng\acute{a}/ \longrightarrow /m\grave{a}-b\acute{a}ng\acute{a}/ \\ /m\grave{a}-/ + /d\acute{a}nd\acute{a}/ \longrightarrow /m\grave{a}-d\acute{a}nd\acute{a}/ \text{ etc.} \end{array}$$

Then the second rule, changing the implosive into an explosive in the environment of the plural morpheme, applies:

Rule 15.

$$\left[\begin{array}{l} -\text{cont} \\ +\text{voice} \\ -\text{breath} \end{array} \right] \longrightarrow [+expl.] \quad / \quad \# \quad \{ma\} \quad \uparrow -V... \quad \left. \vphantom{\begin{array}{l} -\text{cont} \\ +\text{voice} \\ -\text{breath} \end{array}} \right]_{N5} \quad \text{PLURAL}$$

This rule yields representations with the plain-voiced bilabial [b] and alveolar [d] plosives as the initial segments of the forms:

$$\begin{array}{l} /m\grave{a}b\acute{a}ng\acute{a}/, m\grave{a}b\grave{e}r\grave{e}/ \\ /m\grave{a}d\acute{a}nd\acute{a}/, m\grave{a}d\acute{u}r\grave{a}/ \text{ etc}/ \end{array}$$

But since these two segments do not occur as single onsets to a syllable in the language, the final rule can either devoice them, make them breathy-voiced, or, as we saw with Rule 10, make them implosive since [p, b or ɓ]; [t, d, ɗ], can occur in that environment: it so happens that they are devoiced:

Rule 16.

$$\left[\begin{array}{l} -\text{cont} \\ +\text{voice} \\ -\text{breath} \\ +\text{expl.} \end{array} \right] \longrightarrow [-\text{voice}] \quad / \quad \# \quad \{ma\} \quad \uparrow -V... \quad \left. \vphantom{\begin{array}{l} -\text{cont} \\ +\text{voice} \\ -\text{breath} \\ +\text{expl.} \end{array}} \right]_{N6}$$

This rule produces the correct phonetic representations in square brackets in examples (xiv)(a) and (b) above.

When, however, the class 5 noun begins with one

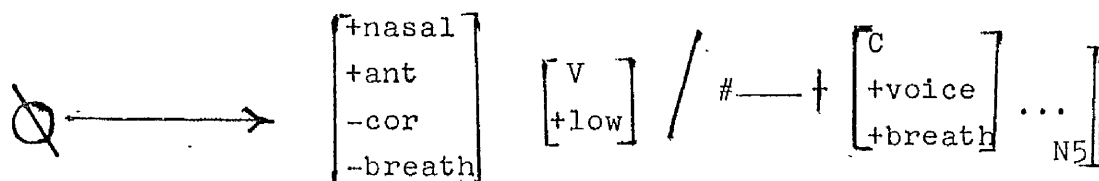
of the following breathy-voiced segments, the devoicing rule does not apply. Instead, the plural or class 6 form is derived by the mere preposition of the plural morpheme {ma} :

(xv)	<u>Class 5</u>	<u>Class 6</u>	<u>English</u>
(a) <u>bilabial stop</u>			
	/pèwùrà/	→ [mà-pèwùrà]	"big axe(s)"
	/bàndítí/	→ [mà-bàndítí]	"bandit" Eng.
	/bìpà/	→ [mà-bìpà]	"robber(s) rapist(s)"
	/bòdòrò/	→ [mà-bòdòrò]	"bottle(s)"
(b) <u>alveolar stop</u>			
	/dàdà/	→ [mà-dàdà]	"duck(s)"
	/dàmù/	→ [mà-dàmù]	"dam(s)"
	/dìbà/	→ [mà-dìbà]	"dip tank(s)"
	/dòròbà/	→ [mà-dòròbà]	"town(s)"
(c) <u>labio-dental fric.</u>			
	/yùd̥zi/	→ [mà-yùd̥zi]	"hair(s)"
	/yírí/	→ [mà-yírí]	"wheel(s)" Eng.
	/yìngá/	→ [mà-yìngá]	"dry clay blocks"
(d) <u>dental fricative</u>			
	/z̥aí/	→ [mà-z̥aí]	"egg(s)"
	/z̥ámù/	→ [mà-z̥ámù]	"breast(s)"
	/z̥uʷá/	→ [mà-z̥uʷá]	"day(s)/ sun(s)"
(e) <u>palato-alveolar fric.</u>			
	/ʒ̥ènd̥ʒ̥e/	→ [mà-ʒ̥ènd̥ʒ̥e]	"huge flame(s) of fire"
	/ʒ̥ùmʷí/	→ [mà-ʒ̥ùmʷí]	"fruit name"
(f) <u>alveolar fricative</u>			
	/z̥éré/	→ [mà-z̥éré]	"young maize stock(s)"
	/z̥ìmbìgʷá/	→ [mà-z̥ìmbìgʷá]	"name of fruit"

	<u>Class 5</u>	<u>Class 6</u>	<u>English</u>	
(g) <u>glottal fricative</u>				
	/hùrè/ →	[mà-hùrè]	"harlot(s)"	
	/hàtʃi/ →	[mà-hàtʃi]	"horse(s)"	
	/hànisi/ →	[mà-hànisi]	"harness(es)" < Eng.	
(h) <u>velar stop</u>				<u>Percentage</u>
	/gàbá/ →	[mà-gàbá]	"tin(s)"	} 66%
	/gàdí/ →	[mà-gàdí]	"solid blocks of soil"	
	/gàmbíngà/ →	[mà-gàmbíngà]	"crocodile(s)"	
	/gákè/ →	[mà-gákè]	"cucumber(s)"	

The rule accounting for all these examples is:

Rule 17.



PL.

Although this rule accounts satisfactorily for all the examples (xv) above, it cannot serve as the general rule whereby class 6 nouns are derived from class 5 nouns commencing with breathy-voiced consonants. The examples below show that some nouns commencing with breathy-voiced segments devoice those segments when the plural morpheme is prefixed to them:

(xvi)	<u>Class 5</u>	<u>Class 6</u>	<u>English</u>	<u>%age</u>
	/gùdò/ →	[mà-kùdò]	"baboon(s)"	} 34%
	/gòmò/ →	[mà-kòmò]	"mountain(s)"	
	/gòróngà/ →	[mà-kòróngà]	"gourge(s)"	
	/gùmbò/ →	[mà-kùmbò]	"leg(s)"	

This behaviour is paralleled by that of some breathy-voiced affricates ranging in place of articulation from the labio-dental to the palato-alveolar, namely /bʏ, dʒ, dʒ, dʒ/. A number of class 5 nouns with these affricates as their initial segment undergo devoicing when the plural morpheme is affixed, while others do not.

There being no obvious way of deciding which set of examples to regard as exceptional to the rule of plural form derivation, a word count was made to determine the number of words belonging to each group. The words counted are those listed in Hannan's (1974) Standard Shona Dictionary, hence the percentage score against each representative word list.

Below, each affricate has two lists, (a) and (b): one comprising a sample of those examples which undergo devoicing, and the other those which remain breathy-voiced:

(xvii) (a) /bʏ/ → [pʃ]

<u>Class 5</u>	<u>Class 6</u>	<u>English</u>	<u>Percentage</u>
/bʏpà/ →	[mà-pfupà]	bone(s)	6%
/bʏnì/ →	[mà-pfèni]	baboon(s)	
/bʏndè/ →	[mà-pfunde]	stock of sorghum	
/bʏrà/ →	[mà-pfura]	fruit type	

(b) /bʏ/ → [bʏ]

/bʏí/ →	[mà-bʏí]	"knee(s)"	94%
/bʏáʋísì/ →	[mà-bʏáʋísì]	"wet piece(s) of firewood"	
/bʏorò/ →	[mà-bʏorò]	"type of mouse/mice"	
/bʏunzò/ →	[mà-bʏunzò]	"examination(s)"	

(xviii) (a) /dz/ → [ts]

Class 5	Class 6	English	Percentage
/dzánzà/ →	[mà-tsánzà]	"old blanket(s)"	7%
/dzàtsí/ →	[mà-tsàtsí]	"bundle(s)"	
/dzìmbà/ →	[mà-tsìmbà]	"foot print(s)"	

(b) /dz/ → [dz]

/dzàtsàtsà/ →	[mà-dzàtsàtsà]	"bowel(s) of hare"	93%
/dzìngídzì/ →	[mà-dzìngídzì]	"bee-fly/flyies"	
/dzìnzà/ →	[mà-dzìnzà]	"line of descent"	
/dzìrà/ →	[mà-dzìrà]	"fibre in sweet potato or meat".	

(xix) (a) [dz] → [tʂ]

/dzàtí/ →	[mà-tʂàtí]	"small measure of liquid"	20%
/dzàtsàtsà/ →	[mà-tʂàtsàtsà]	"spider(s)"	
/dzòtí/ →	[mà-tʂòtí]	"large measure of liquid"	

(b) /dz/ → [dz]

/dzìnú/ →	[mà-dzìnú]	"lizard(s)"	80%
/dzìtí/ →	[mà-dzìtí]	"invader(s)"	
/dzàdzí/ →	[mà-dzàdzí]	"scar(s)"	

(xx) (a) /dʒ/ → [tʃ]

/dʒàk ^w í/ →	[mà-tʃàk ^w í]	"surface water and path"	13%
/dʒékà/ →	[mà-tʃékà]	"sharp bladed grass(es)"	
/dʒàkàwàjà/ →	[mà-tʃàkàwàjà]	"jumble"	

(b) /ɖ/ → [ɖ]

<u>Class 5</u>	<u>Class 6</u>	<u>English</u>	<u>Percentage</u>
/ɖɔŋg ^{wé} / →	[mà-ɖɔŋg ^{wé}]	"cock(s)"	87%
/ɖɛré/ →	[mà-ɖɛré]	"gaol(s)"	
/ɖùrú/ →	[mà-ɖùrú]	"ant(s)"	
/ɖájá/ →	[mà-ɖájá]	"strong your chap(s)"	

The percentages of the devoiced sounds compared to those of their consistently breathy-voiced counterparts show that the affixation of the plural (or class 6) prefix to class 5 nouns causes very few breathy-voiced sounds to be devoiced. Such examples may therefore be treated as exceptions. So it does not matter whether a sound is bilabial, alveolar, palato-alveolar or glottal, whether it is a fricative or a plosive: what is important is that a sound is breathy-voiced. By far the greatest majority of breathy-voiced sounds at the beginning of class 5 nouns remain breathy-voiced when the plural prefix is affixed to them: they are all accounted for by rule 17 above.

The question as to what to do with those breathy-voiced segments that do become unvoiced when the plural morpheme is affixed remains. The theory has a mechanism called a "rule feature" (Chomsky and Halle, 1968, pp.172-177; 373-380) for handling items that behave exceptionally to a general rule. Linked to this concept are two types of features: exception features and morphological features. The principal difference between these two types of features is that whereas the fact that particular lexical items are exceptional to a certain rule gives no indication, in terms

of exception features, as to which alternative rule(s) apply to such items, groups of lexical items characterised by morphological features are subject to a particular alternative rule. Thus if all the items that undergo devoicing when the plural morpheme is affixed are characterised by the rule feature Rule 18.

[- Rule 17],

it is obvious that they are subject to the devoicing rule.

Rule 19.

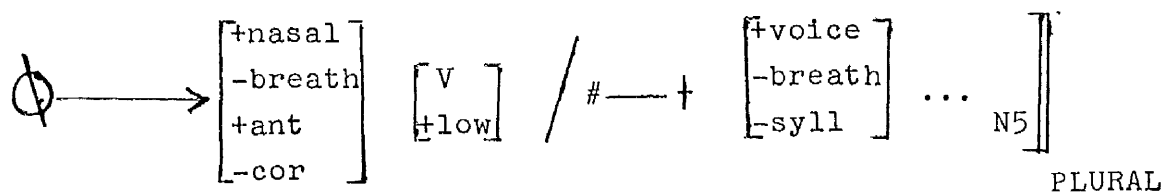
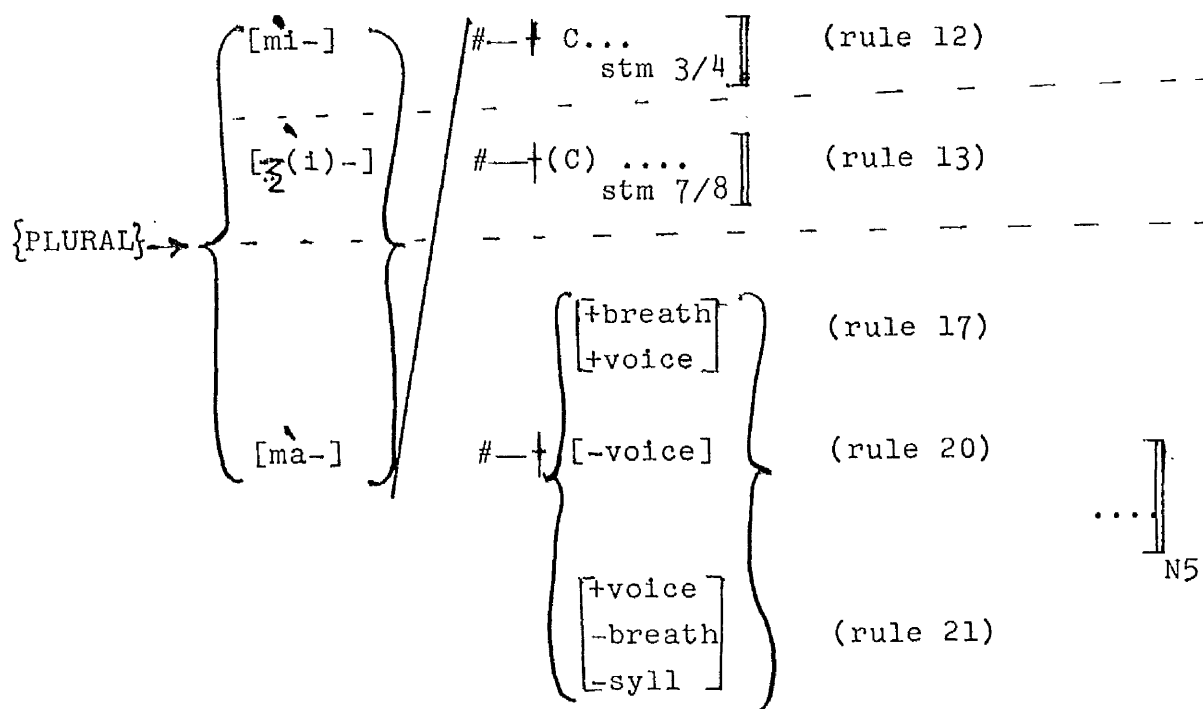
$$\begin{bmatrix} + \text{voice} \\ + \text{breath} \end{bmatrix} \rightarrow \begin{bmatrix} \text{voice} \end{bmatrix} \quad / \quad \# \quad \begin{bmatrix} + \text{nasal} \\ - \text{breath} \\ + \text{ant} \\ - \text{cor} \end{bmatrix} \quad \begin{bmatrix} \text{V} \\ + \text{low} \end{bmatrix} \quad + \text{---} \dots \quad \# \quad \begin{matrix} \text{N5} \\ \text{PLURA} \end{matrix}$$

Rule 19 will not be included in the grammar, since this would be redundant in a system with the morphological rule feature 18.

There are many class 5 nouns whose initial segments are unvoiced, be they fricatives, plosives or affricates. These segments remain unvoiced when the class 6 prefix is affixed to them:

(xxi)	<u>Class 5</u>	<u>Class 6</u>	<u>English</u>
(a)	/fasitèrə/	→ [mà-fasitèrə]	"window(s)" ^{ne(s)}
	/fàsiterà/	→ [mà-fàsiterà]	"window(s)"
(b)	/sàdʒà/	→ [mà-sàdʒà]	"maize meal(s)"
	/sírírí/	→ [mà-sírírí]	"edible caterpillar(s)"
	/sìmbá/	→ [mà-sìmbá]	"power(s)"
(c)	/şòşé/	→ [mà-şòşé]	"ant(s)"
	/şíngá/	→ [mà-şíngá]	"bundle(s) of firewood"
	/şìkìrò/	→ [mà-sìkìrò]	"medium(s) of tutelary spirit"
	/şìngò/	→ [mà-şìngò]	"stone wall(s)"

<u>Class 5</u>	<u>Class 6</u>	<u>English</u>
/ʒòndé/ →	[mà-ʒòndé]	"fig(s)"
/ʒùʒù/ →	[mà-ʒùʒù]	"blind fly/flyes"
(xxiii) /wàrà/ →	[mà-wàrà]	"temerity"
/wéré/ →	[mà-wéré]	"precipice(s)"

Rule 21Rule 22

CHAPTER 7

7.0 SOME TONAL PROCESSES OF KARANGA

7.1.0 Introduction.

The aim of this chapter is mainly to discuss a limited number of tonal processes by way of illustrating how a more comprehensive study of tonal phenomena may be conducted within the theory of transformational generative grammar. The chapter will include the following sections:

1. the definition of a tone language;
2. the relationship between the consonantal margin and syllable pitch;
3. tonal assimilation and dissimilation;
4. the effect of the prefixes of similarity, association and instrumentality on the underlying tone pattern of the nouns of various classes;
5. the stabilization construction and tone, and
6. the object infix and tone.

7.1.1 A tone language.

Doke's (1931) discussion of tone in Shona contains some general remarks on the nature of tone languages. His discussion, like most non-generative phonological studies, is confined to surface tonal phenomena. Doke makes a distinction between what he calls "characteristic" and "significant" tone. Characteristic tone "is the particular method of grouping or succession of musical pitches which characterises a particular language, language group or

language family," (1931, p.215). For example, says Doke, English has a well-defined system of tone, "in which there is generally a rise in pitch on emphasised (stressed) syllables and a fall at the end of sentences."

Significant tone, on the other hand, "plays an active part in the grammatical significance of the language, may be a means of distinguishing words of different meanings otherwise phonetically alike, and may be used to convey varying emotions." Under significant tone, Doke distinguishes three types:

- (a) semantic tone;
- (b) grammatical tone;
- (c) emotional tone.

Semantic tone is what is found in a language such as Chinese, the Cantonese form of which, says Doke, "has no less than seven significant tones."

Grammatical tone is found in the Sudanic languages and Bantu, he says. A tone inflection is sometimes the only factor governing the change in grammatical significance. But, as Doke later demonstrates, Bantu languages have both grammatical and semantic (what I shall call "lexical") tone. Grammatical tone is discussed at length in the last three sections of this chapter: but the following examples may help to show what Doke means by grammatical and semantic tone (see also 6.1.2 above):

Grammatical tone:

/mùʃa/, "a home" vs. /múʃa/, "it's a home"
 /rùkòvǎ/, "a stream" vs. /rúkóvǎ/, "it's a stream",

where the distinction between a nominative form and a

stabilization construction is indicated by tonal substitutions.

Semantic tone:

/kùtʃèrà/, "to dig" vs. /kùtʃérá/, "to fetch water"

/vàngà/, "scar" vs. /vàngá/, "those who were...",

where words with the same segmental phonemes are distinguished in meaning by differences in their tone patterns.

Emotional tone "is much more varied and much more difficult to classify. All emotions of interrogation, sarcasm, emphasis, surprise, doubt, irritation, exultation, command and anger, etc., may be conveyed by a modification of tone," (1931, p. 216). It may be queried whether an interrogative sentence or a command, for example, are emotional sentences. My view is that they are not.

Doke has, however, identified the two most important features of a tone language which are found in Karanga (and Shona as a whole): the use of grammatical and lexical tone. He recognises three tones in Shona, High, Mid and Low. But he says, "From a careful comparison of examples it is evident that of these three tones only two are really significant, that is, relative high and relative low tone." Both in the traditional phoneme theory, and in terms of the generative phonology paradigm, there is no point in recognising three tones for a language if only two of them "are really significant." We are interested, at the classificatory level, only in those tones which are phonologically contrastive. Environmentally conditioned variants (the mid "tone" in Karanga is a Low between two Highs) are redundant and are specified only in phonetic

representations.

However, what Doke (1931, p.217) says about the pitch relationships among the tones of a language is important:

"And it must be emphasised that these tones are never absolute pitches, but only relative, differing in actual pitch from speaker to speaker, from man to woman, from adult to child, and with the same speaker from time to time according to the amount of vigour, fatigue, dullness or brightness which is his at the time of speaking."¹

According to Pike (1948, p.3)² a tone language is any language "having lexically significant, contrastive but relative pitch on each syllable." Whereas significant pitch distinguishes the meanings of utterances (Doke's "grammatical tone"), lexically significant pitch distinguishes the meanings of words, (Doke's "semantic tone"). Pitch contrasts "entail one pitch being kept different from another in the immediate context. Only the relative height of tonemes, not their actual pitch, is relevant to linguistic analysis of tone languages."

7.1.2 Relative [+HIGH] and Relative [-HIGH] tone.

I have said a lot above, in an unstructured way, about the tones of a language being either high or low relative to one another. But the questions how high is [+HIGH] and how low is [-HIGH] need to be probed further. What is the relationship between the lexical categorizations [+HIGH], [-HIGH] on the one hand, and the intonational

-
1. I shall return to the question of relative pitch shortly.
 2. Pike, K.L. (1948) Tone Languages. Ann Arbor, Univ. of Michigan Press.

diversity we perceive?

The main proposal in two recent studies by scholars who have addressed themselves to this question is linked to the assumption that there is a "neutral" median pitch. Ian Maddieson (1970)³ in particular speaks of this median pitch as being some extension of the neutral position of the vocal tract as described by Chomsky and Halle (1968,p.300). This leads him to the view that we should not expect languages with a greater number of level tones to subdivide the same range as languages with fewer tones into smaller intervals, but rather to use a greater range. Pike (1948,p.6) disagrees with this view: "A language with two registers tends to have the contrastive levels farther apart than are the levels of four-register systems." And his illustrative diagrams have the same range for two- three-, and four-toneme systems.

Maddieson maintains, however, that even at the phonetic level we need only deal with pitch as a relative not an absolute, phenomenon. This consideration may have deterred linguists from positing anything as absolute - sounding as a point of origin from which to measure variations of pitch except where this was taken to be the highest or the lowest. Maddieson argues that besides there being phonological justification for the establishment of an unmarked tone level, there is also some evidence of a

3. Maddieson, I. (1970) "The Inventory of Features required for handling Tone in Transformational Phonology", in Tone in Generative Phonology: Research Notes Vol.3, Parts 2 and 3. Department of Linguistics and Nigerian Languages, Univ. of Ibadan.

phonetic nature that the middle of the pitch range of the normal speaking voice is the origin for higher or lower pitches. This was suggested by Longacre⁴ as far back as 1952 in connection with his study of Trique. He posited five phonemic pitch levels for that language, remarking (1952).

"... this brief impressionistic sketch of the phonetic characteristics of the five levels seems to indicate the possibility that the levels with their subphonemic variations could be described in terms of a hypothetical norm of the speaking voice and the degree of departure from the norm... The tonemic oppositions may be conceived of as consisting essentially in a bi-directional spread from a relative norm. This centrifugal opposition is actualised in five relative levels. The centre level, lying close to the phonetic norm, may be considered the structural norm of the system. The highest level represents one pole of the bi-directional tendency, the lowest level represents the other pole."

Maddieson takes this evidence from Trique as confirmation of the existence of

"a linguistic phonetic (not physiological phonetic) pitch norm which is the middle of the voice range in use at any particular time. This norm cannot be given an absolute value in physical terms even for one individual speaker but is at all times relative."

A similar view was expressed recently in a paper that is more interesting because it is a study of tone within the generative phonology model. G.N. Clements (1976)⁵

4. Longacre, R.E. (1952) "Five Phonemic Pitch Levels in Trique". Acta Linguistica, 7, pp.62-8.

5. Clements, G.N. (1976) "Tones as speech melody: a new theory of Terracing," in Papers from the 6th Meeting of the North Eastern Linguistic Society, McGill University. May, Volume 6, pp.49-66.

claims that the relation of these pitch levels to a hypothetical norm is part of the native speaker's knowledge of the language in question. He hypothesises that "the tonal representations generated by the tonal phonology are projected upon sets of abstract tone levels, separately defined for each tone... language. Accordingly for each well-defined tone, there is an associated tone level." These levels, Clements claims, constituting scales, provide a constant frame of reference for the production and recognition of tone melodies,

"abstracting the invariant pattern from the pitch variation due to intonational influences or, (what is particularly relevant to Karanga "depressor" tonology), to segmental conditioning... This frame of reference or scale is assumed to constitute one component of the speaker's full tonal representation."

The scale itself is not to be understood as constituting an absolute set of acoustic parameters. Rather it is subject to modification as a result of the intonational processes that apply to it.

"The identity of the scale itself, however, is not affected by these modifications. In this respect the scale might be compared to a grid drawn upon a flexible sheet, which retains its identity in spite of the distortions which result if the sheet is stretched or twisted."

In other words, the relationship between the scale and the multiplicity of pitches is maintained for each tone in the system, irrespective of the different modes of speech and environment.

In conclusion, Clements says the scale is an ordered set of tone levels, each of which is separated from its neighbour by a (nonnull) interval. If a language has

three well-defined tones, Clements would provide it with a scale consisting of three tone levels, H, M, and L.

"The independently-generated tone formulae are related to the scale by a convention P which projects each tone upon its corresponding tone level. The well-formedness of the relation between the two is maintained through a simple condition requiring each tone to be associated with one and only one tone level throughout the course of a derivation."

Longacre (1952) was concerned mainly to divide the pitch range into two, with a hypothetical neutral above which tone levels are said to be high, and below which they are low. Clements (1976) takes this for granted. He concentrates instead, on providing a mechanism for keeping the various tone levels (both in the high and the low ranges) separate in a multi-tone system such as Longacre's Tique. Their views are complementary.

The concept of a hypothetical pitch neutral from which all other pitches associated with various tones centrifugally diverge, is attractive. The two tones of Shona [+HIGH] and [-HIGH], are realised as a variety of pitches depending on the phonetic environment. And this way of viewing the phenomena brings order to what would otherwise be a confusing situation.

7.1.3 Downstep and Downdrift.

Downdrift is an intonational phenomenon characterising some tone languages, including Shona. It is the automatic lowering in pitch of the second of two high tones between which there is a low tone in the same

utterance. Downdrift is one of the features of the structure of language which anyone learning a tone language will have to know.

The difference between downdrift and downstep was very clearly set out in The Principles of the International Phonetic Association.⁶ In that publication, Igbo is described in the following terms:

"The essential tones are (1), high level... (2) low level... and (3) a lowered high tone shown by ◀ preceding the syllable. All high tones following ◀ are on a slightly lower level than those preceding until the end of the sense-group, when fully high tones are generally resumed. The lowerings shown by ◀ have grammatical or other significance,"

i.e., this is downstep. It is not automatic.

"High tones also undergo a certain lowering whenever a low tone precedes ... lowerings caused in this way likewise continue to the end of the sense-group, after which the normal high of high tones is resumed. These latter lowerings are in the nature of tone assimilation. They have no semantic value, and it is therefore not necessary to mark them in transcriptions."

Except that this neglects to mention that low tones are also lowered, this latter part of the I.P.A. quotation is a good working description of downdrift.

Our understanding of phenomena such as downstep and downdrift depends very much on the adequacy of the phonological theory within which research is being conducted. Some studies are preoccupied with notation and make only the basic phonological distinction between what is and what is not significant. Different theories are designed

6. The Principles of the International Phonetic Association, 1949. (Reprinted, 1972), p.45

to achieve different goals. Sharp (1954)⁷, comparing a monosystemic approach to tonal phonology with a polysystemic one says of the former, that it

"... achieves (as does its sister methodology 'phonemics') what is essentially an overall reading transcription best adapted for the recording of individual utterances or individual texts neither its assumptions nor its findings must be held to have cogency outside their own legitimate sphere."

We cannot condemn a theory for failing to achieve what it was never intended to achieve, seems to be the thrust of Sharp's remarks. The problem arises, I suppose, when excessive claims are made for a theory by its proponents.

The theory espoused in this and the last chapter attempts to explain the surface phenomena in terms of derivation by ordered rules from an underlying simpler structure. Fivaz (1970) says "The reason for the difference between these two treatments (the taxonomic, which forms the main body of his book on Shona Morpho-phonemics and Morphosyntax, and the transformational generative, his Appendix II, pp.224-39) is essentially due to the transformational deep:surface structure distinction as opposed to the surface structure nature of taxonomic analysis."

7. Sharp, A.E. (1954) "A tonal analysis of the disyllabic noun in the Machama dialect of Chaga". Bulletin of the School of Oriental and African Studies, Vol.XVI, No.1: 157-169. Also in Jones and Laver (eds) op.cit., pp.305-319.

7.2.0 The consonantal margin and syllable pitch.

In general, whereas the consonantal margin affects the pitch of a syllable, the reverse is not the case. The consonantal margin plays an important role in determining the pitch characteristics of the two tones of Karanga, [+HIGH] and [-HIGH]. There seems to be an affinity between unvoiced consonants and high pitch on the one hand, and between voiced consonants and low pitch on the other. Of the two types of voiced consonants in Karanga, the plain-voiced and the breathy-voiced, the latter are noted for the lowering effect which they have on the pitch of syllables of which they are the initial segment. This is why these consonants are sometimes called "depressors". A comparison of the pitch of the first syllables of the utterances [mándá] "animal fat" and [mándá] "fork in branch of tree", Mingo-grams 4 and 5 respectively, will show this.

The pitch of the first example is higher and level whereas that of the second starts very low and glides upwards as the nucleus of the syllable is reached, away from the depressor margin:

[mómó] "crowd" [— —]

[zúzúruḍḍa] "drag along the ground" [— — - -]⁸

-
8. Fortune, G. (1968, pp.27-31) contains an extensive study of the various phonetic environments which determine the pitch at which the two tones of the language are realised. I shall therefore not go into detail here: that study being well researched and amply illustrated.

7.2.1 Tonal Assimilation and Dissimilation.

These two processes were briefly discussed in Chapter 1, in the review of Fortune's (1968) Shona Grammatical Constructions. There remains, however, one seemingly puzzling point to raise about them as formulated by Fortune (1968, p.30). He ~~as~~says that assimilation occurs in

- (1) sequences of substantives;
- (2) substantive phrases;
- (3) sequences of verb and object complement, and
- (4) in certain auxiliary verb constructions.

Taking case (3) as an example, we may have the following:

/kùtórá/ "to take" + /sàḍḗ/ "food" [kùtórá sàḍḗ]

/kùbá/ "to steal" + /ḡómbe/ "cattle" [kùbá ḡómbe]

Case (1) may be exemplified by:

/mùrúme/ "a man" + /mùrèfú/ "tall" [mùrúme mùrèfú].

This means that a low tone coming immediately after a high one in the preceding word in the same sense-group, is raised to [+HIGH]:

Rule 23.

/...-H/ → [+H] / § + H # — § ...]
N.A.V.

(where "... " = any tone pattern; "§" = syllable boundary;

"N.A.V." = noun, adjective or verb.)

Thus, assimilation results in a sequence of high tones across word boundaries.

Tonal dissimilation, on the other hand, results in sequences of high and low tones across word boundaries, even though the grammatical constructions may be the same

as those resulting in assimilation allowing [...+H#+H...]:

/mùkáǎzǐ/ "wife" + /wǎŋgù/ "mine" → [mùkáǎzǐ wǎŋgù]

/kùróvǎ/ "to beat" + /nóká/ "snake" → [kùróvǎ nóká]

Rule 24.

i.e. /...+H/ → [-H]/... \$ — # [+H] \$...]

N.A.V.

However, if the two words in one sense-group have low tone next to their boundary, neither assimilation nor dissimilation occurs: whether the construction concerned is a verb-object complement, a sequence of substantives, etc., does not seem to be relevant:

/mùñù/ "a person" + /mùtémá/ "black" → [mùñù mùtémá]

/kùtšàkà/ "to look for" + /sàǎzà/ "food" → [kùtšàkà sàǎzà]

Understanding of tonal assimilation and dissimilation is enhanced by postulating an underlying tonal representation for the words concerned, from which a phonetic representation is derived. Such a representation will differ in shape depending on the syntactic marking of the brackets in phonological surface structures.

7.2.2 Downdrift in Karanga.

This has already been defined as the intonational process whereby the absolute pitch of the second of two [+H] or [-H] syllables is lowered if the syllable between them bears the opposite tone.

The exact value of the amount of lowering does not seem to matter very much, since this will vary from context to context. What is constant is the relationship between [+H] and [-H] in the same sense-group. A low

syllable following a downdrifted high syllable will also be lowered so that their relationship prior to downdrift is maintained.

One way of formalising this is as follows: we shall specify all high tones as $[+H:Pitch_n]$, while all low tones are $[-H:Pitch_m]$: $Pitch_{n/m}$ being the average level for high and low tones respectively. We hypothesize that the pitch drop for downdrifted high tones is the same as that for downdrifted low tones, and give this drop the value 1.

The rule for the first downdrifted $[+H]$ syllable will therefore be:

Rule 25.

$[+H:P_n] \longrightarrow [+H:P_{n-1}]/\#[+H:P_n] \$ [-H:P_m] \$ \longrightarrow$

The second downdrifted high syllable will be $[+H:P_{n-2}]$ and so on. Similarly, the rule for the first downdrifted low syllable will be:

Rule 26.

$[-H:P_m] \longrightarrow [-H:P_{m-1}]/\# [-H:P_m] \$ [+H:P_n] \$ \longrightarrow$

(In both rules "#" marks the beginning of the sense-group and "\$" marks a syllable boundary).

7.3.0 The effect of the adverbial prefixes of "similarity" and "association" on the tone pattern of some nouns.

Our aim in this section is to posit some general rules whereby the underlying tone patterns of various nouns are changed by the affixation of certain grammatical prefixes. The two prefixes which are involved here are the adverbial prefixes of "similarity" and "association", so-

called because they express the idea of "being like Nx" and "with" or "in the company of Nx", respectively.

In section 6.1.2 some ways in which phonology depends on information from the syntactic structure of a string were discussed briefly. In this section we are concerned with those nominal phrases marked

# #	# #	or	# #	# #
sim		sim		assoc.		assoc.

for "similarity construction" and "association construction" respectively in their phonological surface structures.

These adverbial prefixes will have the underlying forms /sà-/ and /nà-/. Although their underlying tone remains the same, i.e. [-HIGH], these prefixes sometimes undergo processes of substitution and coalescence, being realised segmentally as /sò-, sè/ and /nò-, nè/ before the nouns of certain classes.

Because there is so much symmetry between these two adverbial prefixes and since they seem to cause the same tonal perturbations on nouns of similar tone patterns, they were included in the same rules. The way the study was conducted was by grouping the nouns of each class into tone groups, affixing the adverbial prefixes and then observing the resulting tonal modifications.

7.3.1 The nouns of class 1 and 2.

(1)

<u>Tone Ptn</u>	<u>Class 1</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
-H-H	/mù-ṇù/ →	[nò-mú-ṇù]	[sò-mú-ṇù]
	"a person"	with a person	like a person

<u>Tone ptn</u>	<u>Class 1</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
	/mù-rùṅgù/ →	[nò-mù-rùṅgù]	[sò-mù-rùṅgù]
	"a white man"	with a white man	like a white man
	/mù-ròjì/ →	[nò-mù-ròjì]	[sò-mù-ròjì]
	"a witch"	with a witch	like a witch
-H+H+H	/mù-ḡí/ →	[nó-mù-ḡí]	[só-mù-ḡí]
	"an eater"	with an eater	like an eater
	/mù-ṅ ^w í/ →	[nó-mù-ṅ ^w í]	[só-mù-ṅ ^w í]
	"a drinker"	with a drinker	like a drinker
	/mù-rúmé/ →	[nó-mù-rúmé]	[só-mù-rúmé]
	a man	with a man	like a man
	/mù-káḍzì/ →	[nó-mù-káḍzì]	[só-mù-káḍzì]
	a woman	with a woman	like a woman
	/mù-kúró/ →	[nó-mù-kúró]	[só-mù-kúró]
	an older person	with an older person	like an older person
-H+H-H	/mù-ḡórè/ →	[nó-mù-ḡórè]	[só-mù-ḡórè]
	a guy	with a guy	like a guy
	/mù-tóḡ ^w à/ →	[nó-mù-tóḡ ^w à]	[só-mù-tóḡ ^w à]
	an outsider	with an outsider	like an outsider
	/mù-kóma/ →	[nó-mù-kóma]	[só-mù-kóma]
	older brother	with the older brother	like the older brother
	mù-rámu/ →	[nó-mù-rámu]	[só-mù-rámu]
	sister-in-law	with sister-in-law	like sister-in-law
-H+H-H+H	/mù-kómáná/ →	[nó-mù-kómáná]	[só-mù-kómáná]
	a boy	with a boy	like a boy
	/mù-síkàná/ →	[nó-mù-síkàná]	[só-mù-síkàná]
	a girl	with a girl	like a girl

<u>Tone ptttn</u>	<u>Class 1</u>	<u>with /nà-/</u>	<u>with/sà-/</u>
	/mù-rùmbɔ̀àná/ →	[nó-mù-rùmbɔ̀àná]	[só-mù-rùmbɔ̀àná]
	a male heir	with a male heir	like a male heir
-H-H+H	/mù-ròmbó/ →	[nò-mù-ròmbó]	[sò-mù-ròmbó]
	a pauper	with a pauper	like a pauper
	/mù-ròzì/ →	[nò-mù-ròzì]	[sò-mù-ròzì]
	a member of the	with	like a
	Rozvi clan		
-H+H+H-H	/mù-pópótì/ →	[nó-mù-pópótì]	[só-mù-pópótì]
	an announcer	with an announcer	like an announcer
	/mù-núgúnà/ →	[nó-mù-núgúnà]	[só-mù-núgúnà]
	a young brother	with a younger brother	like a younger brother
	/mù-kúwáfà/ →	[nó-mù-kúwáfà]	[só-mù-kúwáfà]
	a brother-in-law	with a brother-in-law	like a brother-in-law
	/mù-k ^w égurù/ →	[nó-mù-k ^w égurù]	[só-mù-k ^w égurù]
	an elderly person	with an elderly person	like an elderly person
	/mù-ʃakábyu / →	[nó-mù-ʃakábyu]	[só-mù-ʃakábyu]
	the deceased	with the deceased	like the deceased

The tonal processes caused by these two adverbial prefixes follow a general pattern: the prefixes, with underlying [-HIGH] tone, assimilate to the tone of the initial syllable of the noun stem. And, just as these prefixes consistently assume the same tone as the initial syllable of the noun stem, so the noun class prefix (with underlying [-HIGH] tone) consistently assumes the opposite tone to that of the initial syllable. Since, however,

these prefixes have different vowels depending on the class to which the noun to which they are affixed belongs, in writing the rules we shall use the abstract forms [Sim.Prefix] and [Assoc.Prefix], for "similarity" and "association" prefix respectively. So the rule accounting for all the examples above will have the form:

Rule 27.

$$\left[\begin{array}{l} \text{-HIGH} \\ \text{SIM or ASSOC.} \\ \text{prefix} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{-HIGH} \\ \text{N}_1\text{-Prefix} \end{array} \right] + \left[\begin{array}{l} \text{-HIGH} \\ \text{...} \end{array} \right]$$

Sim/Ass

(where "α" = "the same as"; "N₁-Prefix" = "the prefix of noun class 1").

The same rule will account for the nouns of class 2, with the following modifications: (1) the N₂-prefix is /v̀à-/, instead of /m̀ù-/: (2) the prefixes of similarity and association will retain their underlying segmental form /s̀à-/ and /ǹà-/. So the rule can be made more general as follows:

Rule 28.

$$\left[\begin{array}{l} \text{-HIGH} \\ \text{SIM or ASSOC.} \\ \text{prefix} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{-HIGH} \\ \text{N}_{1/2}\text{-Prefix} \end{array} \right] + \left[\begin{array}{l} \text{-HIGH} \\ \text{...} \end{array} \right]$$

Sim/Ass.

7.3.2 The nouns of class 1a.

The nouns of this class have no explicit prefix. So in the following examples, the prefixes of association and similarity are affixed to the noun stems directly.

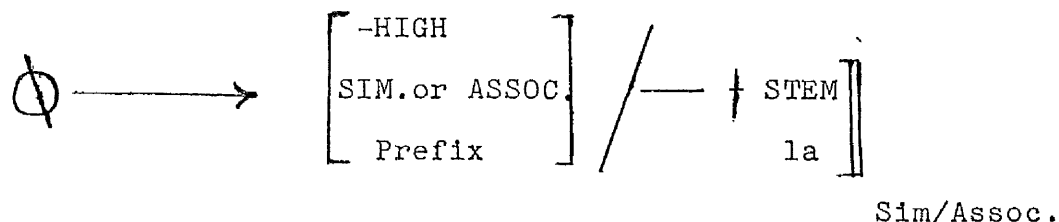
(11)

<u>Tone Ptn</u>	<u>Class 1a.</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
-H+H	/n̩ngí/ →	[nà-n̩ngí]	[sà-n̩ngí]
	so-and-so	with so-and-so	like so-and-so
	/sísí/ →	[nà-sísí]	[sà-sísí]
	sister	with sister	like sister
	/ngàná/ →	[nà-ngàná]	[sà-ngàná]
	so-and-so	with so-and-so	like so-and-so
	/bàbá/ →	[nà-bàbá]	[sà-bàbá]
	father	with father	like father
+H	/ʃé/ →	[nà-ʃé]	[sà-ʃé]
	the chief	with the chief	like the chief
+H+H-H	/tézàrà/ →	[nà-tézàrà]	[sà-tézàrà]
	father-in-law	with father-in-law	like father-in-law
	/sá ^w írá/ →	[nà-sá ^w írá]	[sà-sá ^w írá]
	a ritual friend	with a ritual friend	like a ritual friend
+H+H+H	/sékurú/ →	[nà-sékurú]	[sà-sékurú]
	uncle	with uncle	like uncle
+H-H+H	/néwàndz'é/ → Chief's favourite son.	[nà-néwàndz'é] with the	[sà-néwàndz'é] like.....
+H-H	/ténzì/ →	[nà-ténzì]	[sà-ténzì]
	an owner	with an owner	like an owner
-H-H+H	/tʃìrèmbá/ →	[nà-tʃìrèmbá]	[sà-tʃìrèmbá]
	doctor	with doctor	like doctor

The affixation of either the similarity or the association prefix to the nouns of class 1a of whatever tone pattern does not seem to lead to any change in the

underlying tone pattern: nor is the underlying tone of the prefixes themselves changed. So the rule for the derivation of either the similarity or association construction seems to be fairly straightforward:

Rule 29.



7.3.3 The nouns of class 2a.

When the prefix /vǎ-/ appears before the stems of certain nouns, it indicates the honorific plural and such nouns (with the prefix /vǎ-/) belong to class 2a. The prefixes /mǎḍẓi-/ and /vǎnǎ-/ on the other hand, indicate the numerical plural of class 1a nouns which, by virtue of those prefixes, belong to noun class 2a. The prefixes of similarity and association behave differently in these two types of context:

(iii)

<u>Tone Ptn</u>	<u>Class 2a</u>	<u>with /nǎ-/</u>	<u>with /sǎ-/</u>
(1) a.			
+H-H	/vǎ-ʃè/	→ [nǎ-vǎʃè]	[sǎ-vǎʃè]
	the chief	with the chief	like the chief
	/vǎ-tèzàrà/	→ [nǎ-vǎ-tèzàrà]	[sǎ-vǎ-tèzàrà]
	the father-in-law	with the father-	like the father-
		in-law	in-law
	/vǎ-sèkùrù/	→ [nǎ-vǎ-sèkùrù]	[sǎ-vǎ-sèkùrù]
	the uncle	with the uncle	like the uncle

Tone Ptttn	Class 2a	with /ná-/	with /sà-/
+H-H+H	/vá-tèté/ →	[ná-vá-tèté]	[sà-vá-tèté]
	the aunt	with the aunt	like the aunt
	/vá-m ^w ene/ →	[ná-vá-m ^w ene]	[sà-vá-m ^w ene]
	the mother-	with the mother	like the mother
	in-law	-in-law	-in-law
b. +H+H-H	/vá-sádža/ →	[ná-vá-sádža]	[sà-vá-sádža]
	Mr. Sadza	with Mr Sadza	like Mr Sadza
	/vá-šuro/ (cl. 9/10) →	[ná-vá-šuro]	[sà-vá-šuro]
	Mr Shuro	with Mr Shuro	like Mr Shuro
(2) -H-H+H	/màdžì-mái/ →	[ná-màdžì-mái]	[sà-màdžì-mái]
	the mothers	with the mothers	like the mothers
	/màdžì-šé/ →	[ná-màdžì-šé]	[sà-màdžì-šé]
	the chiefs	with the chiefs	like the chiefs
	/màdžì-sékúru/ →	[ná-màdžì-sékúru]	[sà-màdžì-sékúru]
	the uncles	with the uncles	like the uncles
	/màdžì-mbújà/ →	[ná-màdžì-mbújà]	[sà-màdžì-mbújà]
	the Grannies	with the Grannies	like the Grannies
-H-H-H	/màdžì-tàtégùru/ →	[ná-màdžì-tàtégùru]	[sà-màdžì-tàtégùru]
	the Ancestors	with the Ancestors	Like the Ancestors
	/màdžì-šàbà/ →	[ná-màdžì-šàbà]	[sà-màdžì-šàbà]
	the Fathers	with the Fathers	like the Fathers
	/màdžì-màmbo/ →	[ná-màdžì-màmbo]	[sà-màdžì-màmbo]
	the Kings	with the Kings	like the Kings
(3)			
-HIH-H+H	/vána-šòšé/	[ná-vána-šòšé]	[sà-vána-šòšé]
	Mr Svosve and Co.	with Mr Svosve	like Mr Svosve
		and co.	and co.
	/vána-ningí/	[ná-vána-ningí]	[sà-vána-ningí]
	So-and-so and co	with so-and-so and co	Like so-and-so and co.

Class 2a examples seem to fall into three major categories in as far as their response to the affixation of the two adverbial prefixes is concerned. Those in category (1) can be accounted for by the following rule:

Rule 30.

$$\emptyset \longrightarrow \left[\begin{array}{c} \text{-HIGH} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] / \text{A} + \left[\begin{array}{c} \text{+HIGH} \\ \text{N}_{2a}\text{-Prefix} \end{array} \right] + \text{STEM} \quad \left. \vphantom{\left[\begin{array}{c} \text{-HIGH} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right]} \right] \quad \text{Sim/Assoc.}$$

The difference between the examples (1) a. and b. is that the latter are proper names ordinarily designating objects other than human. They have been adapted from their usual classes (indicated), whereas group a. are intrinsic 2a examples. The division into group 1a. and b. has no tonal implications, however.

When the adverbial prefixes of similarity and association are affixed to the nouns in category (2), with the class prefix /mād̥zi-/ , i.e. /-H-H/,

- (a) the tone pattern of the class 2a stem is not changed, but
- (b) the first syllable of the class prefix becomes [+H], while
- (c) the second remains [-H];
- (d) the adverbial prefixes retain their underlying tones:

Rule 31.

$$\left[\begin{array}{c} \text{-HIGH} \\ \text{mād̥zi} \end{array} \right] \longrightarrow [+H-H] / \left[\begin{array}{c} \text{-HIGH} \\ \text{SIM or Assoc.} \\ \text{Prefix} \end{array} \right] + \text{---} + \text{STEM} \quad \left. \vphantom{\left[\begin{array}{c} \text{-HIGH} \\ \text{SIM or Assoc.} \\ \text{Prefix} \end{array} \right]} \right] \quad \text{Sim/Assoc.}$$

With the examples in category (3), both the prefix and the 2a stem retain their lexical tone pattern. But the adverbial prefixes of similarity and association are raised from their lexical [-HIGH] to the phonetic [+HIGH]:

Rule 32.

$$\left[\begin{array}{l} \text{-HIGH} \\ \text{Sim Or Assoc} \\ \text{Prefix} \end{array} \right] \rightarrow \left[\text{+HIGH} \right] / \text{---} \dagger [\text{v}^{\text{à}}\text{n}^{\text{á}}] \dagger \text{STEM} \quad \left. \vphantom{\left[\begin{array}{l} \text{-HIGH} \\ \text{Sim Or Assoc} \\ \text{Prefix} \end{array} \right]} \right] \text{Sim/Assoc.}$$

7.3.4 The nouns of classes 3 and 4.

When affixed to the nouns of class 3, the adverbial prefixes of similarity and association, with underlying forms /sà-/ and /nà-/, are realised as /so-/ and /no-/ in Karanga, their tone depending on factors to be discussed below. And since the nouns of class 3 have the prefix /mù-/, the rule governing their combination with these adverbial prefixes is predictably the same as that posited for the nouns of class 1 above, which also have the prefix /mù-/.

(iv)

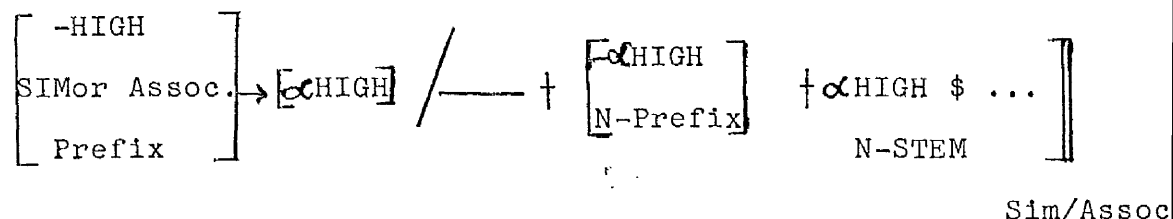
<u>Tone Ptn</u>	<u>Class 3</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
-H-H...	/mù-nàmàtò/	→ [nò-mú-nàmàtò]	[sò-mú-nàmàtò]
	a prayer	with a prayer	like a prayer
	/mù-rìḍ̣̀ẓ̣̀i/	→ [nò-mú-rìḍ̣̀ẓ̣̀i]	[sò-mú-rìḍ̣̀ẓ̣̀i]
	instrumentalist	with an instru-	like an instru-
		mentalist	mentalist
	/mù-ḍ̣̀ẓ̣̀i/	→ [nò-mú-ḍ̣̀ẓ̣̀i]	[sò-mú-ḍ̣̀ẓ̣̀i]
	a root	with a root	like a root

<u>Tone Ptn</u>	<u>Class 3</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
	/mù-ṇṇgùrú/ →	[nò-mú-ṇṇgùrú]	[sò-mú-ṇṇgùrú]
	a " " tree	with a " " tree	like a " " tree
-H-H+H..	/mù-tàmbá/ →	[nò-mú-tàmbá]	[sò-mú-tàmbá]
	a " " tree	with a " " tree	like a " " tree
	/mù-ṣàmbá/ →	[nò-mú-ṣàmbá]	[sò-mú-ṣàmbá]
	a " " plant	with a " " plant	like a " " plant
	/mù-rìrìró/ →	[nò-mú-rìrìró]	[sò-mú-rìrìró]
	a characteristic	with a character-	like a character-
	ring	istic ring	istic ring
-H+H..	/mù-ṣá/ →	[nó-mù-ṣá]	[só-mù-ṣá]
	a home	with a home	like a home
	/mù-púrúró/ →	[nó-mù-púrúró]	[só-mù-púrúró]
	ulutation	with ulutation	like ulutation
	/mù-sóró/ →	[nó-mù-sóró]	[só-mù-sóró]
	a head	with a head	like a head
	/mù-sána/ →	[nó-mù-sána]	[só-mù-sána]
	a back	with a back	like a back
	/mù-kámbo/ →	[nó-mù-kámbo]	[só-mù-kámbo]
	a factory	reject with a factory	like a factory
		reject	reject

These examples show that the same rule that accounts for the tonal processes resulting from the combination of the two adverbial prefixes with the nouns of classes 1 and 2 also operates here. So rule 28 need not be so specific as to pinpoint classes 1 and 2 in the environment since classes 3 and 4 are also covered. That rule will now have to be modified so that it accounts not only for the examples from classes 1, 2, 3 and 4 (the class

prefix of 4, the numerical plural of the nouns of class 3, is /mì-/), but also for most Karanga nouns with a class prefix both in the underlying and in the derived representation; all nouns, that is, except those of classes 1a, 2a, 5, 9 and 10:

Rule 33.



This rule states that in a similarity or association phrase, the prefixes of similarity and association will always have the same tone as the initial syllable of a noun stem, itself preceded by its noun class prefix, bearing the opposite tone.

7.3.5 The nouns of classes 5, 9 and 10.

As we have said in a previous section (6.3.2), the class prefix of nouns in class 5 has been lost. The tonal processes brought about by the affixation of the two adverbial prefixes /nà-/ and /sà-/ (realised with classes 5, 9 and 10 as /ne-/ and /se-/) are different from those caused by the same prefixes when attached to nouns whose class prefix is extant.

(v)

<u>Tone Ptn</u>	<u>Class 5</u>	<u>with /nà-/</u>	<u>with /sà-/</u>
-H-H	/gùdò/ →	[nè-gùdò]	[sè-gùdò]
(1)	a baboon	with a baboon	like a baboon
	/gùmbò/ →	[nè-gùmbò]	[sè-gùmbò]
	a leg	with a leg	like a leg

Tone Ptn	Class 5	with /nà-/	with /sà-/
	/gùròkùrò/ →	[nè-gùròkùrò]	[sè-gùròkùrò]
	a gullet	with a gullet	like a gullet
	/zè ^v èzè ^v è/ →	[nè-zè ^v èzè ^v è]	[sè-zè ^v èzè ^v è]
	whispering	with whispering	like whispering
+H-H	/dúm ^w à/ →	[né-dúm ^w à]	[sé-dúm ^w à]
	a charm	with a charm	like a charm
(2)	/dób ^y è/ →	[né-dób ^y è]	[sé-dób ^y è]
	wild fruit	with wild fruit	like wild fruit
	/tsimè/ →	[né-tsimè]	[sé-tsimè]
	a well	with a well	like a well
+H+H	/gùrúm ^w andirá/ →	[né-gùrúm ^w andirá]	[sé-gùrúm ^w andirá]
	a big crowd	with a big crowd	like a big crowd
	/bàrikà/ →	[né-bàrikà]	[sé-bàrikà]
	bigamy	with bigamy	like bigamy
	/dèndèrà/ →	[né-dèndèrà]	[sé-dèndèrà]
	a ground hornbill	with a ground hornbill	like a ground hornbill
-H+H	/gàbà/ →	[né-gàbà]	[sé-gàbà]
	a tin	with a tin	like a tin
(3)	/gòróngà/ →	[né-gòróngà]	[sé-gòróngà]
	a ditch	with a ditch	like a ditch
	/dèng ^w aná/ →	[né-dèng ^w aná]	[sé-dèng ^w aná]
	a basket	with a basket	like a basket
	/gèhèná/ →	[né-gèhèná]	[sé-gèhèná]
	hell	with hell	like hell

Tone Ptn	Classes 9 and 10 with /nà-/	with /sà-/
(1)	-H-H... /hòʃà/ → [nè-hòʃà]	[sè-hòʃà]
	an illness	with an illness like an illness
	/ŋòmbe/ → [nè-ŋòmbe]	[sè-ŋòmbe]
	cattle	with cattle like cattle
	/ʃòkà/ → [nè-ʃòkà]	[sè-ʃòkà]
	feet	with feet like feet
	/hàràbɔ̃ʔa/ → [nè-hàràbɔ̃ʔa]	[sè-hàràbɔ̃ʔa]
(2)	an old man	with an old man like an old man
	/hùrùdʒa/ → [nè-hùrùdʒa]	[sè-hùrùdʒa]
	a successful farmer	with a successful farmer like a successful farmer
	+H+H... /hùrùʔa/ → [né-hùrùʔa]	[sé-hùrùʔa]
	dust	with dust like dust
	/hàngíro/ → [né-hàngíro]	[sé-hàngíro]
	a frying pan	with a frying pan like a frying pan
(2)	/maridʒo/ → [né-maridʒo]	[sé-maridʒo]
	a sermon	with a sermon like a sermon
	/mòpòtò/ → [né-mòpòtò]	[sé-mòpòtò]
	a row	with a row like a row
	/hàkàtə/ → [né-hàkàtə]	[sé-hàkàtə]
	lots	with lost like lots
	+H-H+H /hàngàŋ ^w a/ → [né-hàngàŋ ^w a]	[sé-hàngàŋ ^w a]
	forgetfulness	with forgetfulness like forgetfulness
	/hàkàná/ → [né-hàkàná]	[sé-hàkàná]
	a small pot	with a small pot like a small pot
	/hùk ^w aná/ → [né-hùk ^w aná]	[sé-hùk ^w aná]
	a chicken	with a chicken like a chicken
	/hòfòrí/ → [né-hòfòrí]	[sé-hòfòrí]
	a stout person	with a stout person like a stout person

<u>Tone Ptn</u>	<u>Classes 9 and 10 with /na-/</u>	<u>with /sa-/</u>
-H+H	/ḥàpxá/	[né-ḥàpxá]
	an armpit	like an armpit
(3)	/ḥàwí/	[né-ḥàwí]
	craving	like craving
	/džùngú/	[né-džùngú]
	peanuts	like peanuts
	/ḥùrití/	[né-ḥùrití]
	a green pigeon	like a green pigeon

What is common to the examples (1) in both class 5 and classes 9 and 10 above, is that the first of a sequence of [-HIGH] tones, or the one [-HIGH] syllable in a monosyllabic stem, is raised to [+HIGH] in the phonetic representation after the affixation of the adverbial prefixes: the latter retain their underlying [-HIGH] tone:

Rule 34.

$$[-H(-H...)] \xrightarrow{N5,9,10} [+H(-H...)] / \left[\begin{array}{l} \text{-HIGH} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] + \text{---} \parallel$$

Sim/Assoc.

In other words, there is dissimilation: both the adverbial prefixes and the first syllable of the noun have underlying [-HIGH], but the latter is raised when the two are combined.

Examples (2) have in common that any high tone(s) at the beginning of a class 5,9 or 10 stem are lowered when a similarity or association prefix is affixed: but in any particular stem, a high tone coming after a low one remains high. In all cases, the adverbial

prefixes are raised from their underlying [-HIGH] to [+HIGH]. So there are two rules: first the adverbial prefixes assimilate to the tone of the first syllable of the noun stem:

Rule 35.

$$\left[\begin{array}{l} \text{[-HIGH]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] \longrightarrow \left[\begin{array}{l} \text{[+HIGH]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] / \text{---} + \text{---} \left[\begin{array}{l} \text{[+HIGH..]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] \quad \text{N5,9,10.}$$

This rule results in such forms as

/né-dúm^wá/
 /né-déndéra/
 /né-húrú^wá/
 /né-hàngà^wá/

Such forms are not acceptable: there must be tonal polarization across the boundaries in these constructions. So the high tones of the noun stem (except those coming after an intervening low tone) are lowered:

Rule 36.

$$[+H(+H -H+H)] \longrightarrow [-H(-H-H+H)] / \left[\begin{array}{l} \text{[+HIGH]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] + \text{---} \left[\begin{array}{l} \text{[+HIGH]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] \quad \text{Sim/Assoc.}$$

The polarization in tone realization which is a feature of the above examples can be seen at work in the examples (3). These examples have an initial [-HIGH] followed by a [+HIGH] syllable. First the [-HIGH] adverbial prefix is affixed to the noun stem.

Rule 37.

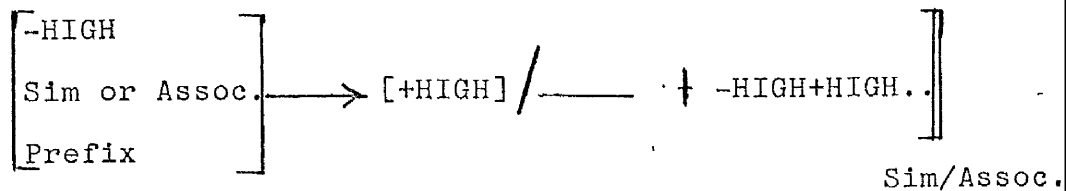
$$\emptyset \longrightarrow \left[\begin{array}{l} \text{[-HIGH]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] / \text{---} + \text{---} \left[\begin{array}{l} \text{[-H+H..]} \\ \text{Sim or Assoc.} \\ \text{Prefix} \end{array} \right] \quad \text{N5,9,10} \quad \text{Sim/Assoc.}$$

This rule gives the following unacceptable forms:

/nè-gàbá/ /sè-gàbá/
 /nè-gòrónḡà/ /sè-gòrónḡà/
 /nè-ḡàpxá/ /sè-ḡàpxá/, unacceptable because

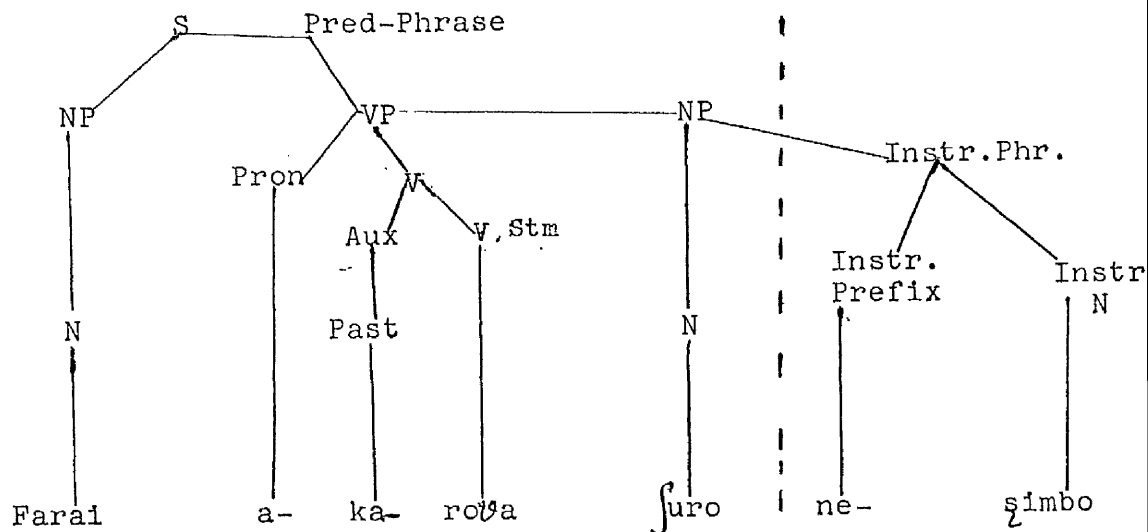
they have a sequence of like tones across a morpheme boundary. So, as in the examples (2), the adverbial prefixes are raised to [+HIGH], while, here, the noun stem retains its underlying tone pattern:

Rule 38.



7.4.0 The Adverbial prefix of instrumentality.

In this discussion, it will be assumed that the adverbial prefix of instrumentality collocates with nouns which have the semantic feature [-HUMAN]. The syntagm in which the prefix occurs has the following shape: e.g. [Farái ákàróḡa ʃurò néḡsimbò] "Farai hit a rabbit with a knobkerry."



In this discussion we are interested only in the tonal structure of the instrumentality phrase on the extreme right of the above tree.

The adverbial prefix of instrumentality has the underlying representation /nè-/, i.e. "with" or "using".. When affixed to the nouns of class 3, this prefix sometimes has the shape /nò-/, for some Karanga speakers, e.g.

[nò-mùsèʔe] "with an arrow"

It will have become obvious by now that this prefix is homophonous with some realizations of the prefix of association in the previous section. Although in the context of the nouns of classes 3,4,5,9,10 these two prefixes have the same phonological shape, they are

semantically distinct, as their English renditions will show. But this distinction may not be [mʷàná wàʔámásám̐bà ásiyàna ánoḡonàtʃíkó] where the two prefixes combine with the same nouns.

For example, we have an instrumentality and an association construction using the word /muti/ "medicine":

- (a) [nó-mù-tí] "with medicine": association.
- (b) [nó-mù-tí] "using medicine": instrumentality.

In sentences:

- (a) [n[ndàsijá sádʒà rákó nómùtí pɛʒo nɛʃotó]

"I have left your food together with your medicine near the hearth".

- (b) [mʷàná wàʔámásám̐bà ásiyàna ánoḡonàtʃíkóró nómùtí]

"Mr Masambaasijana's⁹ son excels in class using medicine".

Because the adverbial prefix of instrumentality

9. This is a popular name given to wily medicine-men or witch-doctors.

has the same underlying tone as the prefixes of association and similarity, there is no need to posit some rules governing the tonal processes resulting from its combination with the nouns of the various classes in Karanga. The same rules whereby these prefixes were seen to combine with the nouns of classes 3,4,5,9,10... will also serve here. And those five classes were chosen only as examples. The adverbial prefix of instrumentality does not combine with some nouns, e.g. those belonging to classes 1,2,1a and 2a.

7.4.1 The Stabilization Construction.

By the "stabilization construction" we mean essentially the same structure as defined by H. Carter (1956)¹⁰, i.e. "a form which is capable of standing by itself as a complete sentence, other than a free verbal tense". There are many word-types that can be stabilized, e.g. independent and dependent nominals, proper names. But in this discussion we shall limit ourselves to independent nominals belonging to a number of noun classes. We shall also limit our discussion to affirmative stabilizations, i.e. those translated

"it's..."

and leave negative ones, i.e. those translated

"it isn't..."

Both types are discussed by Carter (1956) for Manyika.

10. Carter, Hazel (1956) "Stabilization in the Manyika Dialect of the Shona Group" Africa, XXVI,4,pp.398-405.

The stabilization constructions we have chosen fall into three groups

- (i) those indicated by a high-tone class prefix;
- (ii) those indicated by the stabilizing element /i'-/;
- (iii) those indicated by the stabilizing element /ndi'-/.

7.4.1.1 Stabilization Constructions with a class prefix.

There are many independent nominals with class prefixes and which belong to many different noun classes. As such, it would be cumbersome to treat each of the classes separately. We shall therefore use examples from classes 1 and 2, 3 and 4, 6, and 7 and 8, and posit some rules which will predict the way other nouns from the remaining prefixed classes will behave.

(vi)

<u>Tone Ptn</u>	<u>Class 1</u>	<u>Stabilized form</u>	<u>English</u>
(-H)-H	/mù-ṇù/ →	[mu'-ṇù]	he's/it's a person a person
(1)	/mù-ràndà/ →	[mu'-rándà]	he's/it's a subject a subject
	/mù-rùṅgu/ →	[mu'-rùṅgu]	he's/it's a whiteman a whiteman
(-H)+H-H+H	/mù-síkàná/ →	[mu'-síkàná]	(s)he's/it's a girl a girl
(2)	/mù-kòmàná/ →	[mu'-kòmàná]	he/it's a boy a boy
	/mù-rùṁbyàná/ →	[mu'-rùṁbyàná]	he's a male heir a male heir

<u>Tone Pattn</u>	<u>Class 1</u>	<u>Stabilized form</u>	<u>English</u>
(-H)+H+H..	/mù-rùmé/ →	[mú-rùmé]	he's a man a man
(3)	/mù-kádží/ →	[mú-kádží]	she's a woman a woman
	/mù-róóorà/ →	[mú-ròòorà]	she's a bride a bride
	/mù-kómbórerì/ →	[mú-kòmbòrerì]	he's a benefactor a benefactor
(-H)-H+H	/mù-ròmbó/ →	[mú-ròmbó]	he's a pauper a pauper
(4)	/mù-tṣené/ →	[mú-tṣené]	he's a holy one a holy one
<u>Class 3</u>			
(-H)-H-H	/mù-nàmàto/ →	[mú-nàmàto]	it's a prayer a prayer
(1)	/mù-rídží/ →	[mú-rídží]	he's a instrumentalist an instrumentalist
	/mù-nùngùrú/ →	[mú-nùngùrú]	it's a mùngùrú tree a " " tree
(-H)+H-H+H	/mù-ṣòzòwá/ →	[mú-ṣòzòwá]	it's a mu o owa tree a " " tree
(2)	/mù-tṣìndzìk ^w á/ →	[mú-tṣìndzìk ^w á]	it's a cross a cross
(-H)+H+H	/mù-purúru/ →	[mú-purúru]	it's ululation ululation
(3)	/mù-sóro/ →	[mú-sòrò]	it's a head a head

<u>Tone Ptn</u>	<u>Class 3</u>	<u>stabilized form</u>	<u>English</u>
	/mù-ḡédḡédḡé/	→ [mú-ḡédḡédḡé]	it's a hair style a hair style
(-H)-H+H	/mù-tàmbá/	→ [mú-tàmbá]	it's a mùtàm bá tree a " " tree
(4)	/mù-ḡàmbá/	→ [mú-ḡàmbá]	it's a mùḡàmbá plant a " " plant
	/mù-ḡòmbó/	→ [mú-ḡòmbó]	it's snuff mixture snuff mixture
	/mù-rèrèró/	→ [mú-rèrèró]	it's a way of bringing way of bringing up child up a child
<u>Class 6</u>			
(-H)-H-H	/mà-kùdò/	→ [má-kùdò]	they are baboons baboons
(1)	/mà-kùmbò/	→ [má-kùmbò]	they are legs legs
	/mà-fàḡaḡaḡaḡa/	→ [má-fàḡaḡaḡaḡa]	it's abundance abundance
	/mà-kùròkùrò/	→ [má-kùròkùrò]	they are gullets gullets
(-H)+H-H+H	/mà-púrìsá/	→ [má-púrìsá]	they are policemen policemen
(2)	/mà-ḡùḡùḡ ^w á/	→ [má-ḡùḡùḡ ^w á]	they are sloughs sloughs
	/mà-tsikombí/	→ [má-tsikombí]	they are big girls big girls
	/mà-kùkùḡ ^w á/	→ [má-kùkùḡ ^w á]	it's debris. debris

<u>Tone Ptn</u>	<u>Class 6</u>	<u>Stabilized form</u>	<u>English</u>
(-H)+H+H..	/mà-kòkò/ →	[má-kòkò]	they are bits of crust bits of crust
(3)	/mà-kótó/ →	[má-kòtò]	it's chaff chaff
	/mà-kòmbòrèrò/ →	[má-kòmbòrèrò]	they are blessings blessings
(-H)-H+H..	/mà-sìmbá/ →	[má-sìmbá]	they are powers powers
(4)	/mà-pùràzí/ →	[má-pùràzí]	they are farms farms
	/mà-tìkítì/ →	[má-tìkítì]	they are tickets Eng. tickets
	/mà-pùràŋga/ →	[má-pùràŋga]	they are planks Eng. planks

Class 7

(-H)-H-H	/tʃí-rùŋgù/ →	[tʃí-rùŋgù]	it's English English
(1)	/tʃí-ñù/ →	[tʃí-ñù]	it's a thing a thing
	/tʃí-ŋá/ →	[tʃí-ŋá]	it's malice malice
	/tʃí-rùŋgùrìrà/ →	[tʃí-rùŋgùrìrà]	it's heartburn heartburn
(-H)+H-H	/tʃí-ḡag ^w é/ →	[tʃí-ḡag ^w é]	it's maize maize
(2)	/tʃí-ḡukù/ →	[tʃí-ḡukù]	it's a small book a small book
	/tʃí-ndig ^w aná/ →	[tʃí-ndig ^w aná]	it's a small plate a small plate

<u>Tone Ptttn</u>	<u>Class 7</u>	<u>Stabilized form</u>	<u>English</u>
	/tʃí- mbúdzàná/ → [tʃí- mbúdzàná]	it's a kid	
	a kid		
	/tʃí- bókìsì/ → [tʃí- bókìsì]	it's a small box	Eng.
	a small box		
<hr/>			
(-H)+H+H..	/tʃí- g ^w índírí/ → [tʃí- g ^w índírí]	he's a short strong	
	a short strong person	person	
(3)	/tʃí- pápátá/ → [tʃí- pápátá]	it's bread	
	bread		
	/tʃí- ɓatíró/ → [tʃí- ɓatíró]	it's a handle	
	a handle		
<hr/>			
(-H)-H+H..	/tʃí- zòndó/ → [tʃí- zòndó]	it's an ankle	
	an ankle		
(4)	/tʃí- gǎnú/ → [tʃí- gǎnú]	it's a milk container	
	milk container		
	/tʃí- tàvúnà/ → [tʃí- tàvúnà]	it's a small Kaross	
	a small Kaross		

Examples from classes 4 and 8 have been omitted from the above lists because they differ from those of classes 3 and 7 (respectively), their numerical singulars, only in the segmental shape of their class prefixes - /mù-/ ~ /mì-/ for 3 and 4; and /tʃí(1)-/ ~ /z(1)-/ for 7 and 8. Otherwise the stabilization construction is derived using the same rules.

In the lists above, the examples from each class are divided into four categories, depending on their tone pattern. And rules of the stabilized form derivation will be posited for each category. What is common to all these

examples is that the class prefix, with an underlying [-HIGH] tone, is raised to [+HIGH] in the stabilization construction:

Rule 39.

$$\left[\begin{array}{c} \text{-HIGH} \\ \text{N-Prefix} \end{array} \right] \longrightarrow \text{[+HIGH]} \quad / \text{---} \text{+} \quad \text{N-STEM} \quad \left. \vphantom{\left[\begin{array}{c} \text{-HIGH} \\ \text{N-Prefix} \end{array} \right]} \right] \text{STAB.}$$

(where "STAB" = Stabilization construction).

The rest of this discussion will be designed to systematise the repercussions for the tone pattern of the noun stem of raising the class prefix to [+HIGH].

All examples marked (1) above have stems commencing with a [-HIGH] syllable, which is followed either by zero or by another [- HIGH] syllable. The tone of the third syllable, if there is a third syllable, does not affect the process. Only the first [-HIGH] syllable is raised to [+HIGH] in the stabilized form:

Rule 40.

$$\left[\text{-H(-H, ...)} \right] \xrightarrow{\text{N-STEM}} \left[\text{+H(-H, ...)} \right] \quad / \quad \left[\begin{array}{c} \text{+HIGH} \\ \text{N-Prefix} \end{array} \right] \quad \text{+---} \quad \left. \vphantom{\left[\begin{array}{c} \text{+HIGH} \\ \text{N-Prefix} \end{array} \right]} \right] \text{STAB}$$

So here we have a tonal assimilation across 'morpheme' boundaries, where, in the similarity, association and instrumentality constructions we have a polarization of tone realizations. Whether we have assimilation or polarization, therefore, is grammatically, rather than phonologically, determined. So these are transformational phonological rules, as opposed to non-transformational ones, which do not take the syntactic categorization of an item into account.

All the examples marked (2) have in common that

the class prefix and the initial syllable of the stem are on opposite lexical tones: [-HIGH] and [+HIGH] respectively. This dissimilation is maintained in the stabilization construction. Given the general rule 39: that the class prefix is raised in this construction, this means that the initial syllable of the stem is lowered. The second feature shared by the examples (2) is that their second syllable is [-HIGH]. And any [+HIGH] syllable coming after that second syllable, is unaffected by the stabilization derivation:

Rule 41.

$$[+H-H..(+H)] \xrightarrow{\text{N-stem}} [-H-H..(+H)] / \left[\begin{array}{c} +\text{HIGH} \\ \text{N-Prefix} \end{array} \right] + \text{---} \parallel \text{Stab}$$

Examples (3) have stems commencing with sequences of at least two [+HIGH] tones. In the stabilization construction, only the first two tones are lowered, unless a sequence of three [+HIGH] tones is followed by a [-HIGH] fourth syllable, in which case all three [+HIGH] are lowered:

Rule 42.

$$(a) \quad [+H+H(+H)] \xrightarrow{\text{N-STEM}} [-H-H(+H)] / \left[\begin{array}{c} +\text{HIGH} \\ \text{N-Prefix} \end{array} \right] + \text{---} \parallel \text{STAB}$$

$$(b) \quad [+H+H+H-H] \xrightarrow{\quad} [-H_4] \quad / \quad \left[\begin{array}{c} +\text{HIGH} \\ \text{N-Prefix} \end{array} \right] + \text{---} \parallel \text{STAB}$$

("H₄" is an ad hoc abbreviation for [-H-H-H-H])

All the stems in group (4) commence with a [-HIGH] syllable immediately followed by one that is [+HIGH].

This tone pattern is maintained in the stabilization construction:

Rule 43.

$$\begin{bmatrix} \text{-HIGH} \\ \text{N-Prefix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{+HIGH} \end{bmatrix} / \text{---} + \begin{bmatrix} \text{N-STEM} \\ \text{-H+H} \end{bmatrix}$$

STAB

7.4.1.2 Stabilization constructions indicated by /i'-/

To illustrate this construction, only the nouns of classes 5 and 9/10 will be used. These have also been grouped into tonal categories.

(vii)

<u>Tone Ptn</u>	<u>Class 5</u>	<u>Stabilized form</u>	<u>English</u>
-H-H..	/gùdò/ → a baboon	[i'-gùdò]	it's a baboon
(1)	/gùmbò/ → a leg	[i'-gùmbò]	it's a leg
	/gùròkùrò/ → a gullet	[i'-gùròkùrò]	it's a gullet
	/zèzèzèzè/ → whispering	[i'-zèzèzèzè]	it's whispering
+H-H	/dùm ^w à/ → a charm	[i'-dùm ^w à]	it's a charm
(2)	/dòb ^y è/ → wild fruit	[i'-dòb ^y è]	it's a wild fruit
	/tsìme/ → a well	[i'-tsìme]	it's a well
	/gùrù ^w andirà/ → a crowd	[i'-gùrù ^w andirà]	it's a crowd

<u>Tone Ptn</u>	<u>Class 5</u>	<u>Stabilized form</u>	<u>English</u>
	/ḡárikà/ →	[i'-ḡárikà]	it's bigamy bigamy
	/ḡèndèrà/ →	[i'-ḡèndèrà]	it's a ground horn- ground hornbill bill
-H+H..	/ḡàbá/ →	[i'-ḡàbá]	it's a tin a tin
(3)	/ḡòróṅga/ →	[i'-ḡòróṅga]	it's a ditch a ditch
	/ḡèṅg ^w aná/ →	[i'-ḡèṅg ^w aná]	it's a basket a basket
	/ḡèḡhèná/ →	[i'-ḡèḡhèná]	it's hell hell
(4)+H-H	/ḡíjǎ/ →	[i'-ḡíjǎ]	it's a gear a gear
	/ríṽa/ →	[i'-ríṽa]	it's a trap a trap
<u>Class 9/10</u>			
-H-H..	/ḡòṣa/ →	[i'-ḡòṣa]	it's an illness an illness
(1)	/ḡòmbe/ →	[i'-ḡòmbe]	they are cattle cattle
	/ṣòkà/ →	[i'-ṣòkà]	they are feet feet
	/ḡàràbǎ/ →	[i'-ḡàràbǎ]	it's an old man an old man
	/ḡùrùḡza/ →	[i'-ḡùrùḡza]	it's a successful a successful farmer farmer
	/ḡ ^w ajaná/ →	[i'-ḡ ^w ajaná]	it's a lamb a lamb
+H+H-H	/ḡùrùṽa/ →	[i'-ḡùrùṽa]	it's dust dust

<u>Tone ptttn</u>	<u>Class 9/10</u>	<u>stabilized form</u>	<u>English</u>
	/hàngíró/ →	[i'-hàngíró]	it's a frying pan a frying pan
(2)	/marídzo/ →	[i'-marídzo]	it's a sermon a sermon
	/mópótò/ →	[i'-mópótò]	it's a row a row
	/hákàtà/ →	[i'-hákàtà]	they are lots lots
-H+H	/hàpxá/ →	[i'-hàpxá]	it's an armpit an armpit
	/hàwí/ →	[i'-hàwí]	it's a craving craving
(3)	/dzùngú/ →	[i'-dzùngú]	it's a pea-nut pea-nut
	/hùritì/ →	[i'-hùritì]	it's a green pigeon a green pigeon
+H-H	/hùm ^w e/ →	[i'-hùm ^w e]	it's a beer working- beer working-party party
(4)	/hàḡàná/ →	[i'-hàḡàná]	it's a small pot a small pot
	/nàritì/ →	[i'-nàritì]	it's a needle a needle
	/fòḡorò/ →	[i'-fòḡorò]	it's a shovel a shovel

What is common to all these examples is the
general rule:

Rule 44. $\Phi \longrightarrow [i'] / \text{---} \uparrow$ N5,9/10
STEM

||
STAB.

All the examples marked (1) are made up of stems commencing with at least two [-HIGH] tones (although there are [-HIGH] monosyllables in both classes). The initial syllable is raised to [+HIGH] when the stabilization prefix [i'-] is affixed to these nouns:

Rule 45.

$$[-H(-H...)] \longrightarrow [+H(-H...)] / [i'-] \text{ + } \text{---} \text{] } \begin{array}{c} \text{N5,9/10} \\ \text{STAB} \end{array}$$

At least the initial syllable of the examples marked (2) and (4) is on a [+HIGH] tone. In the environment of the stabilization prefix this first syllable (or the first two high syllables) is lowered to [-HIGH]: if there is a third high syllable, it remains high: being lowered only when followed by a low: The examples (2) and (4) here are therefore similar to the examples (3) in section 7.4.1.1 above and can be accounted for by the rules 42a and b, with the stabilization prefix [i'-] in the place of the $\begin{bmatrix} +\text{HIGH} \\ \text{N-Prefix} \end{bmatrix}$:

Rule 46.

$$(a) \quad [+H(+H \ +H)] \longrightarrow [-H(-H \ +H)] / [i'-] \text{ + } \text{---} \text{] } \begin{array}{c} \text{N5,9/10} \\ \text{STAB} \end{array}$$

$$(b) \quad [+H+H+H-H] \longrightarrow [-H_4] / [i'-] \text{ + } \text{---} \text{] } \begin{array}{c} \text{N5,9/10} \\ \text{STAB} \end{array}$$

All the examples marked (3) are made up of stems whose first two syllables are on [-HIGH] and [+HIGH] tone respectively. This tone pattern remains unchanged in the stabilization construction, which is derived by affixing [i'-] to the stem:

Rule 47.

$$\emptyset \longrightarrow [i'] / \text{---} + \begin{bmatrix} -H + H \\ N5,9/10 \end{bmatrix}$$

STAB

7.4.1.3 Stabilization constructions indicated by /ndí'-/

Independent nominals which take the prefix /ndí'-/ in the stabilized construction belong to noun classes 1a and 2a.

(viii)

<u>Tone Ptnn</u>	<u>Class 1a</u>	<u>Stabilized form</u>	<u>English</u>
-H-H..	/mámbo/ →	[ndí'-mámbo]	he's a king a king
(1)	/hòmbàrùme/ →	[ndí'-hòmbàrùme]	he's a successful successful hunter hunter
	/tʃìrembá/ →	[ndí'-tʃìrembá]	it's doctor doctor
-H+H	/sìsì/ →	[ndí'-sìsì]	it's sister sister
(2)	/bàbá/ →	[ndí'-bàbá]	it's father father
	/ngàná/ →	[ndí'-ngàná]	it's so-and-so so-and-so
	/nìngí/ →	[ndí'-nìngí]	it's so-and-so so-and-so
+H+H..	/težàrà/ →	[ndí'-težàrà]	it's father-in-law father-in-law
(3)	/sàh ^w íra/ →	[ndí'-sàh ^w íra]	he's a ritual a ritual friend friend
	/sékurú/ →	[ndí'-sékurú]	it's uncle uncle

<u>Tone Ptn</u>	<u>Class 1a</u>	<u>Stabilized form</u>	<u>English</u>
+H-H+H	/ʋá-tè-té/ →	[ndí-ʋá-tè-té]	it's aunt aunt
(4) +H	/ʃe/ →	[ndí-ʃe]	he's a chief a chief

For examples (1) and (4), the initial syllable of the stem is realised on a tone opposite to that which it has in the underlying representation:

Rule 48.

$$[\alpha \text{HIGH}; .] \rightarrow [-\alpha \text{HIGH..}]/[\text{ndí-}] + \text{---} \Bigg] \text{STAB}$$

The majority of the examples above, marked (2) and (3) do not change their tone pattern when the stabilization prefix is affixed to them:

Rule 49.

$$\emptyset \rightarrow [\text{ndí-}]/\text{---} + \text{1a STEM} \Bigg] \text{STAB}$$

(ix)

<u>Tone Ptn</u>	<u>Class 2a</u>	<u>Stabilized form</u>	<u>English</u>
(+H)-H-H	/ʋá-tè-zà-à/ →	[ndí-ʋá-tè-zà-à]	it's the father- the father-in-law in-law
	/ʋá-tèn-zì/ →	[ndí-ʋá-tèn-zì]	he's the owner the owner
(+H)-H+H	/ʋá-m ^w èné/ →	[ndí-ʋá-m ^w èné]	it's the father the mother-in-law in-law
(+H)+H+H-H	/ʋá-sàh ^w í-à/ →	[ndí-ʋá-sàh ^w í-à]	he's the ritual the ritual friend friend
(+H)+H-H	/ʋá-tít-à/ →	[ndí-ʋá-tít-à]	he's the teacher the teacher

Eng.

<u>Tone Ptn</u>	<u>Class 2a</u>	<u>Stabilized form</u>	<u>English</u>
(-H+H)-H+H	/v`aná-n`ngí/ →	[ndí'-v`aná-n`ngí]	they are the
	the so-and-sos		so-and-sos
	/v`aná-şòşé/ →	[ndí'-v`aná-şòşé]	it's Mr Svosve
	Mr Svosve and co.		and co.
(-H+H)+H-H	/v`aná-sad`za/ →	[ndí'-v`aná-sad`za]	it's Mr Sadza
	Mr Sadza and co.		and co.
	/v`aná-şúró/ →	[ndí'-v`aná-şúró]	it's Mr Shuro and
	Mr Shuro and co.		co.
	/v`aná-mámbo/ →	[ndí'-v`aná-mámbo]	they are the kings
	the kings		

Although the examples above fall into several tone categories, and into two groups depending on the prefix, their tone patterns remain unchanged when the stabilization prefix [ndí'-] is affixed to them:

Rule 50.

$$\emptyset \longrightarrow [ndí'-] / \text{---} + \left\{ \begin{array}{l} [v`a-] \\ [v`aná] \end{array} \right\} + \left. \begin{array}{l} \text{2a STEM} \\ \text{STAB} \end{array} \right\}$$

There is another category of class 2a nouns whose prefix is /m`ad`zi-/. These nouns behave differently in two ways: they do not take the stabilization prefix [ndí'-]; they undergo tonal modification when the stabilization construction is derived:

(x)

<u>Tone Ptn</u>	<u>Class 2a</u>	<u>Stabilized form</u>	<u>English</u>
(-H-H)+H..	/m`ad`zi-şé/ →	[m`ad`zi-şé]	they are the chiefs
	the chiefs		
(1)	/m`ad`zi-sékúru/ →	[m`ad`zi-sékúru]	they are the uncles
	the uncles		

<u>Tone Ptn</u>	<u>Class 2a</u>	<u>Stabilized form</u>	<u>English</u>
	/màd̥z̥i-maí/	→ [mád̥z̥i-maí]	they are the mothers the mothers
	/màd̥z̥i-mbújà/	→ [mád̥z̥i-mbújà]	they are the grannies the grannies
<hr/>			
(-H-H)-H+H	/màd̥z̥i-tàtégùrù/	→ [mád̥z̥i-tàtégùrù]	they are the the Ancestors Ancestors
(2)	/màd̥z̥i-ḡaḡa/	→ [mád̥z̥i-ḡaḡa]	they are the fathers the fathers
	/màd̥z̥i-tèté/	→ [mád̥z̥i-tèté]	they are the aunts the aunts

These examples fall into two categories. In the first one, where at least the first syllable bears [+HIGH] tone, the stabilization construction is formed by modifying the tone pattern of the class prefix, that of the stem remaining unchanged.

Rule 51.

$$\left[\begin{array}{c} \text{[-HIGH]} \\ \text{màd̥z̥i} \end{array} \right] \longrightarrow \left[\begin{array}{c} \text{[+H-H]} \\ \text{---} \end{array} \right] + \left[\begin{array}{c} \text{+H...} \\ \text{2a STEM} \end{array} \right]$$

STAB

This rule accounts for all 2a stems above except those whose first two syllables are [-H+H..]. In the latter cases, i.e. stems which are [-H+H], both syllables of the class prefix are raised to [+HIGH]:

Rule 52.

$$\left[\begin{array}{c} \text{[-HIGH]} \\ \text{màd̥z̥i} \end{array} \right] \longrightarrow \left[\begin{array}{c} \text{[+HIGH]} \\ \text{---} \end{array} \right] + \left[\begin{array}{c} \text{-H+H} \\ \text{2a STEM} \end{array} \right]$$

STAB

Rule 53

<div data-bbox="297 331 561 1608"> <div data-bbox="346 340 512 446">[+HIGH N-Prefix]</div> <div data-bbox="346 723 429 787">[i'-]</div> <div data-bbox="346 1149 479 1212">[ndi'-]</div> <div data-bbox="379 1404 512 1468">[maɖzi]</div> <div data-bbox="379 1532 512 1596">[maɖzi]</div> </div>	<div data-bbox="24 893 297 957">{STAB.}</div>	<div data-bbox="693 127 1371 212"># — + +H(-H)... \parallel_{N STEM} (rule 40)</div>
		<div data-bbox="693 212 1371 297"># — + +H-H...(+H) \parallel_{N-STEM} (rule 41)</div>
		<div data-bbox="693 340 1371 446"># — + { [-H(-H +H)] } \parallel_{N-STEM} (rule 42)</div>
		<div data-bbox="693 446 1371 532"># — + -H+H \parallel_{N-STEM} (rule 43)</div>
		<div data-bbox="693 595 1371 680"># — + +H(-H) \parallel_{N5,9,10} (rule 45)</div>
		<div data-bbox="693 702 1371 829"># — + { [+H(-H+H)...] } \parallel_{N5,9,10} (rule 46)</div>
		<div data-bbox="693 829 1371 915"># — + -H+H... \parallel_{N5,9,10} (rule 47)</div>
		<div data-bbox="693 957 1371 1064"># — + -αHIGH... \parallel_{1a STEM} (rule 48)</div>
		<div data-bbox="693 1085 1371 1191"># — + { [-H+H...] [+H-H...] } \parallel_{1a STEM} (rule 49)</div>
		<div data-bbox="693 1212 1371 1319"># — + { [ʔá-] [ʔána] } +... \parallel_{2a STEM} (rule 50)</div>
		<div data-bbox="693 1361 1371 1468"># — + +H... \parallel_{2a STEM} (Rule 51)</div>
		<div data-bbox="693 1489 1371 1596"># — + -H+H... \parallel_{2a STEM} (rule 52)</div>

7.5.0 The treatment of the object infix in Karanga.

Since it is transitive verbs that collocate with nouns functioning as objects, this final section will not deal with intransitive verbs. In sentences or clauses where the object of a transitive verb is mentioned explicitly, the syntagmatic relationship between the two is

verb - object

If, as we have done throughout, the infinitive inflection of the verb is taken as the underlying representation, we get the following examples in the language:

(xi)	<u>Verb</u>	<u>Object</u>	<u>English</u>
(a)	/Kù-t̚sàkà	ḥúní/	"To to look for firewood"
(b)	/Kù-páráɖza	m ^w aná/	"To calm a baby"
(c)	/Kù-pórá	tsám̐bà/	"to write a letter"

The object infix arises in clauses such as these when the object is not referred to directly, but by its grammatical concord. Examples (a)-(c) will have the form:

(xii) (a) /Kù-dzí-tʂákà/ "to look for them"
 (b) /Kù-mú-náráɖʂá/ "to calm him/her"
 (c) /Kù-jí-nórá/ "to write it"

(i.e. the concord is "infix" between the two parts of the verb).

Example (a) in particular shows that a transitive verb may have its underlying tone pattern modified when an object infix is used. The purpose of this section is to posit some tonal rules whereby that underlying tone pattern of the transitive verb is affected by the use of

the object infix. The influence of the object infix on other inflections of the verb (past and future tense) was also investigated. In these cases, it was assumed that the derivation of the phonetic representation containing the object infix takes place in several stages: the past tense is derived from the infinitive inflection and then the clause or sentence with the object infix is derived after that. We are only interested in the latter process.

7.5.1 The infinitive inflection and the object infix.

The verbs used in this discussion are of two kinds: those whose stems have low tone, and those which have high tone.

7.5.1.1. [-HIGH] STEMS

(xiii)

<u>Verb infinitive</u>	<u>with Obj. infix</u>	<u>category of infix</u>
/kù-rùmbìdžà/ "to praise" →	[kù-mù-rùmbìdžà]	"him", n.class 1
/kù-ròngèdžà/ "to arrange" →	[kù-zí-ròngèdžà]	"them" n.class 8
/kù-nònòtsà/ "to delay" →	[kù-ɔ́-nònòtsà]	"them" n.class 2
/kù-wònèkà/ "to bid farewell" →	[kù-tí-wònèkà]	"us" 2nd person plural pron.
/kù-jàmùrà/ "to help" →	[kù-kú-jàmùrà]	"you" 2nd person sing. pron.
/kù-yùndùrùrà/ "to unravel" →	[kù-dzì-yùndùrùrà]	"them" n.class 10

Throughout, it is assumed that the grammatical concord or class prefix (functioning as the object infix) has an underlying [-HIGH] tone. Since in all the examples

given above the object infix is realised on a [+HIGH] tone, there is a general rule:

Rule 54.

$$\begin{bmatrix} \text{[-HIGH]} \\ \text{O. I.} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{[+HIGH]} \\ \text{Ku-} \end{bmatrix} / \begin{bmatrix} \text{[-HIGH]} \\ \text{STEM} \end{bmatrix} + \text{---} + \begin{bmatrix} \text{[-HIGH]} \\ \text{STEM} \end{bmatrix} \quad \text{V.T.}$$

(i.e. the object infix, with an underlying [-HIGH] tone, becomes [+HIGH] in the context [-HIGH] infinitive particle to the left and [-HIGH]-commencing transitive verb to the right).

The second stage in the derivation is a tonal assimilation rule: the initial syllable of the verb stem becomes [+HIGH], like the object infix before it:

Rule 55.

$$\begin{bmatrix} \text{[-HIGH...]} \\ \text{V.T.} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{[+H-H...]} \\ \text{V.T.} \end{bmatrix} / \begin{bmatrix} \text{[-HIGH]} \\ \text{Ku-} \end{bmatrix} + \begin{bmatrix} \text{[+HIGH]} \\ \text{O.I.} \end{bmatrix} + \text{---}$$

7.5.1.2 [+HIGH] stems.

In all the following examples, a minimum of one, and a maximum of three, syllables is/are on [+HIGH] tone:

(xiv)

<u>Verb infinitive</u>	<u>with object infix</u>	<u>category of infix</u>
/kù-ʃámíʃá/ "to surprise"	→ [kù-ʃá-ʃámíʃá]	"them" 3rd person pron.pl.
/kù-bùgúmúra/ "to hit hard"	→ [kù-ʒí-bùgúmúra]	"them" n.8
/kù-báʔáʔá/ "to fondle"	→ [kù-mú-báʔáʔá]	"her" 3rd person pron. sing.
/kù-yéʒéŋgúra/ "to shake"	→ [kù-kú-yéʒéŋgúra]	"you" 2nd person pron. sing.

/kù-ǵá/ "to eat"	[kù-rí-ǵá] "it" n.5
/kù-ǵá/ "to steal"	[kù-ǵí-ǵá] "it" n.9

In all the examples the tone pattern of the high tone verbs remains unchanged. So the rule is fairly straightforward:

Rule 56.

$$\left[\begin{array}{c} \text{-HIGH} \\ \text{O.I.} \end{array} \right] \longrightarrow \left[\text{+HIGH} \right] / \left[\begin{array}{c} \text{-HIGH} \\ \text{ku-} \end{array} \right] + \text{---} + \left[\begin{array}{c} \text{+HIGH} \\ \text{STEM} \end{array} \right]$$

7.5.2.0 The past tense inflection and the object infix.

Since the derivation of the past tense inflection of transitive verbs from the underlying infinitive is not central to (but is assumed for) this discussion, it has not been included here, but is briefly set out in Appendix II (pp.449-451). Verbs with an underlying (i.e. infinitive representation) shape of [-HIGH...] behave as follows.

7.5.2.1 [-HIGH] stems: with 1st and 2nd Personal Pronouns as subjects.

(xv)	(1)	(2)	(3)	(4)
	<u>Verb infinitive</u>	<u>Past tense</u>	<u>With ob.inf. and 1st and 2nd Person Pronouns as subjects.</u>	<u>Category of infix</u>
	/kù-x ^w ǎǎ/ →	[t-à-ká-x ^w ǎǎ]	[t-à-ká-dzì-x ^w ǎǎ]	we crushed them N.10
	"to crush"			
	/kù-ǵopèrà/ →	[m-à-ká-ǵopèrà]	[m-à-ká-ǵì-ǵopèrà]	you saw fortune teller about it N.8
	"to fortune tell"			
	/kù-rùkà/ →	[nd-à-ká-rùkà]	[nd-à-ká-tǵì-rùkà]	I knitted it N.7
	"to knit"			

/kù-wóm̀bèkà/ → [nd-à-ká-wóm̀bèkà][nd-à-ká-kù-wóm̀bèkà] I
 "to baptise" baptised you 2nd Per. Pron. Sing.

Given the general rule that a low verb stem has its initial syllable raised to [+HIGH] in the past tense when a first or second person (plural and singular) pronoun is in subject position, (Appendix II) the clause with the object infix is derived as follows:
Rule 57.

$$\emptyset \longrightarrow \left[\begin{array}{c} \text{[-HIGH]} \\ \text{[O.I.]} \end{array} \right] / \left[\begin{array}{c} \text{[-HIGH]} \\ \text{[Noun]} \end{array} \right]_{\text{Subj.}} + \left[\begin{array}{c} \text{[+HIGH]} \\ \text{[PAST]} \end{array} \right] + \left[\begin{array}{c} \text{+H-H...} \\ \text{v.STEM} \end{array} \right]_S$$

(where " $\left[\begin{array}{c} \text{[-HIGH]} \\ \text{[Noun]} \end{array} \right]$ " = 1st or 2nd person (sing. or plural) pronouns serving as subject; " PAST " = [-ká], and "S" = sentence.)

7.5.2.2 [-HIGH] stems: with 3rd Personal Pronouns and [-HUMAN] concords as subjects.

When the subject of the sentence is either the 3rd personal pronoun (singular or plural) or the grammatical concord of a [-HUMAN] noun, i.e. the subject is $\left[\begin{array}{c} \text{[+HIGH]} \\ \text{[Noun]} \end{array} \right]_{\text{Subj.}}$, the tone pattern of these same sentences will be as follows:

	Subj.	{PAST}	obj.infix	Verb
	[+HIGH]	[-HIGH]	[+HIGH]	[+HIGH-HIGH..]
	↓	↓	↓	↓
[(v̀)akómáná)	v́á	ká	tj́í	rú kà]

Note that the verb retains its past tense tone pattern whereby the initial syllable is raised to [+HIGH].

If, as we have been doing, the underlying (surface structure) tone of the morpheme {PAST} is regarded as [+HIGH], the rule for the examples with [+HIGH] grammatical concords as subjects is as follows:

Rule 58.

$$(a) \quad \begin{bmatrix} +HIGH \\ \{PAST\} \end{bmatrix} \rightarrow [-HIGH] / \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \text{---} + \begin{bmatrix} -HIGH \\ \text{O.I.} \end{bmatrix} + \begin{bmatrix} +H-H.. \\ \text{V.STEM} \end{bmatrix}_S$$

Then the tone of the object infix is raised to [+HIGH], like that of the initial syllable of the verb stem:

$$(b) \quad \begin{bmatrix} -HIGH \\ \text{O.I.} \end{bmatrix} \rightarrow [+HIGH] / \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \begin{bmatrix} -HIGH \\ \{PAST\} \end{bmatrix} + \text{---} + \begin{bmatrix} +H-H.. \\ \text{V.STEM} \end{bmatrix}_S$$

(In both (a) and (b), "[+HIGH Noun Subj.]" = 3rd person (sing. or plural) pronoun i.e. [á] and [wá], or all [-HUMAN] concords).

7.5.3.0 [+HIGH] stems: with 1st and 2nd Personal Pronouns as subjects.

As is explained in Appendix II, high stem verbs are lowered to [-HIGH] in the past tense if they are associated with first or second person pronouns functioning as subjects. We want now to see what the effect of introducing an object infix is on the tone pattern of such verbs:

(1) <u>Verb infinitive</u>	(2) <u>Past tense</u>	(3) <u>With object infix and 1st and 2nd Personal Pronoun as subjects.</u>	(4) <u>Category of object infix</u>
-------------------------------	--------------------------	---	--

/kù-téngá/ → [nd-à-ká-tèngà] [nd-à-ká-tfì-téngá] I bought it
"to buy" (N.7)

/kù-séwénzèsà/ → [t-à-ká-sèwénzèsà...] [t-à-ká-zì-séwénzèsà] we
 "to use" used them (N.8)
 /kù-byárúrá/ → [w-à-ká-byárùrà...] [w-à-ká-dzì-byárúrá] you
 "to tear" tore them (N.10)
 /kù-rúmúrá/ → [m-à-ká-rùmùrà...] [m-à-ká-mù-rúmúrá] you (pl.)
 "to wean" weaned him 3rd. per. sing.

As can be seen by comparing the first three columns above, the underlying tone of the verbs (1) is [+HIGH] up to the third syllable: but it is lowered to [-HIGH] in a sentence of the form (2):

Subj. + {PAST} + Verb + object (where the object slot is indicated by "...")

When, however, the object infix is introduced, (3) above, the verb stem regains its underlying tone; the {PAST} morpheme retains its [+HIGH] and the object infix remains [-HIGH]:

Rule 59.

$$\emptyset \longrightarrow \begin{bmatrix} \text{[-HIGH]} \\ \text{[O.I.]} \end{bmatrix} / \begin{bmatrix} \text{[-HIGH]} \\ \text{[Noun]} \end{bmatrix}_{\text{Subj.}} + \begin{bmatrix} \text{[+HIGH]} \\ \text{[PAST]} \end{bmatrix} + \text{---} + \begin{bmatrix} \text{+H+H...} \\ \text{V STEM} \end{bmatrix}_S$$

7.5.3.1 [+HIGH] stems: with 3rd Personal Pronouns and [-HUMAN] concords as subjects.

When the subject noun is represented either by the third person (singular or plural) pronoun or any grammatical concord of non-human nouns, the situation is different to that in the previous section:

(xvii)

(1) Verb infinitive	(2) Past tense	(3) With obj. infix and 3rd pers. pron. or [-HUMAN] concord as subject	(4) Category of infix
/kù-téngá/ →	[á-kà-téngá]	[á-kà-t _i -téngá]	"he bought it" "to buy" (Noun cl. 7)
/kù-sévéngzésà/ →	[v-á-kà-sévéngzésà]	[v-á-kà-z _i -sévéngzésà]	"to use" they used them (N.8)
/kù-byárúrá/ →	[r-á-kà-byárúrá]	[r-á-kà-dz _i -byárúrá]	"it tore" them (N.10)
/kù-rúmúrá/ →	[g ^w -á-kà-rúmúrá]	[g ^w -á-kà-dz _i -rúmúrá]	"it weaned" them" (N.10)

i.e. the verb stems retain their underlying tone in (2)

and (3). What changes is the underlying tone of two morphemes: the {PAST} from [+HIGH], assimilating to the underlying tone of the object infix:

Rule 60.

$$(a) \quad \left[\begin{array}{c} +\text{HIGH} \\ \{ \text{PAST} \} \end{array} \right] \longrightarrow [-\text{HIGH}] / \left[\begin{array}{c} +\text{HIGH} \\ \text{Noun} \\ \text{Subj.} \end{array} \right] + \text{---} + \left[\begin{array}{c} -\text{HIGH} \\ \text{O.I.} \end{array} \right] + \left[\begin{array}{c} +\text{H}+\text{H} \\ \text{V.STEM} \end{array} \right]$$

Then the tone of the object infix assimilates to that of the initial syllable of the high stem verb:

$$(b) \quad \left[\begin{array}{c} -\text{HIGH} \\ \text{O.I.} \end{array} \right] \longrightarrow [+ \text{HIGH}] / \left[\begin{array}{c} +\text{HIGH} \\ \text{Noun} \\ \text{Subj.} \end{array} \right] + \left[\begin{array}{c} -\text{HIGH} \\ \text{PAST} \end{array} \right] + \text{---} + \left[\begin{array}{c} +\text{H}+\text{H} \\ \text{V.STEM} \end{array} \right] \text{S}$$

It is essential that rule (a) should precede (b): the {PAST} must first assimilate to the underlying tone of the object infix, [-HIGH], before the latter in its turn assimilates to the tone of the initial syllable of the high tone verb stem.

7.5.4.0 The Future tense inflection¹¹ and the object infix.

This section, like that dealing with the past tense, is divided into two major parts: one for low stem transitive verbs; and the other for high stem ones. Each of those parts is further subdivided into two subparts: for low stem verbs with 1st and 2nd personal pronouns functioning as subjects, and with the 3rd personal pronouns and [-HUMAN] concords functioning as subjects. The same subdivisions are made under the high stem verbs.

7.5.4.1 [-HIGH] stems: with 1st and 2nd personal pronouns as subjects

(xviii)	(1)	(2)	(3)	(4)
	<u>Verb infinitive</u>	<u>Future tense</u>	<u>With object infix and 1st and 2nd personal pronouns as subjects</u>	<u>Category of infix</u>
	/kù-ràmbà/	→ [nd-ì-tʃá-rámbà]	[nd-ì-tʃá- <u>zì</u> -rámbà]	"I shall reject <u>them</u> " (N.8)
	/kù-rùkà/	→ [..ù-tʃá-rùkà]	[ù-tʃá-tʃì-rùkà]	<u>you</u> (sing.) will knit <u>it</u> (N.7)
	/kù-byùmìrà/	→ [m-ù-tʃá-byùmìrà]	[m-ù-tʃá-mù-byùmìrà]	<u>you</u> (pl.) will allow <u>him</u> (3rd per.sing)
	/kù-wòmbèkà/	→ [t-ì-tʃá-wòmbèkà]	[t-ì-tʃá-kù-wòmbèkà]	<u>we</u> shall baptise <u>you</u> (2nd per.sing.)

Here, as with the past tense, the insertion of the low tone object infix between the [+HIGH] tense marker ([-tʃá-] for future) and the [+HIGH] first syllable of the verb is all that is needed. So Rule 57 above will

11. Rules whereby the Future tense inflection is derived from the underlying infinitive are set out in Appendix III, pp. 452-453.

have to be rewritten to make it more general:

Rule 61.

$$\emptyset \longrightarrow \left[\begin{array}{c} [-\text{HIGH}] \\ \text{O.I.} \end{array} \right] / \left[\begin{array}{c} [-\text{HIGH}] \\ \text{Noun} \end{array} \right]_{\text{Subj.}} + \left[\begin{array}{c} [+ \text{HIGH}]^* \\ \{ \text{PAST} \} \\ \{ \text{FUT} \} \end{array} \right] + \text{Verb} \left[\begin{array}{c} [+ \text{H-H...}] \end{array} \right]_s$$

(* This notation is being used to indicate that both options are acceptable.)

7.5.4.2 [-HIGH] stems: with 3rd personal pronouns and
[-HUMAN] concords as subjects.

The third person pronouns (plural and singular) have high underlying tone - as do the concords of [-HUMAN] nouns.

(xix)

(1)	(2)	(3)	(4)
<u>Verb infinitive</u>	<u>Future tense</u>	<u>With 3rd personal pronoun and [-HUMAN] concords as subjects</u>	<u>Category of object infix</u>

/kù-ràmbà/ → [á-tʃá-rámbà] [á-tʃá-zí-rámbà] "he will reject
"to reject" them" (N.8)

/kù-rúkà/ → [ʋ-á-tʃá-rúkà] [ʋ-á-tʃá-tʃí-rúkà] "they will
"to knit" knit it" N.7

/kù-x^wápà/ → [dz-í-tʃá-x^wápà] [dz-í-tʃá-tí-x^wápà] "they will
"to crush" crush us"

2nd pers.plural
/kù-byòvòrà/ → [r-í-tʃá-byòvòrà] [r-í-tʃá-jí-byòvòrà] "it will
"to pierce" pierce it" N.9

The rule operating in these examples, like rule 58 above, has two parts: and the object infix and the future

tense marker undergo similar processes as the comparable morphemes in that rule, which will now be incorporated in a more general one:

Rule 62

$$(a) \quad \begin{bmatrix} +HIGH \\ \{PAST\} \\ \{FUT.\} \end{bmatrix} \rightarrow [-HIGH] / \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \text{---} + \begin{bmatrix} -HIGH \\ \text{O.I.} \end{bmatrix} + \begin{bmatrix} +H-H \\ \text{V.STEM} \end{bmatrix} \quad S$$

i.e. the tense markers assimilate to the tone of the object infix which they immediately precede. The object infix itself then assimilates to the tone of the initial syllable of the inflected verb:

$$(b) \quad \begin{bmatrix} -HIGH \\ \text{O.I.} \end{bmatrix} \rightarrow [+HIGH] / \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \begin{bmatrix} -HIGH \\ \{PAST\} \\ \{FUT.\} \end{bmatrix} + \text{---} + \begin{bmatrix} +H-H \\ \text{V.STEM} \end{bmatrix} \quad S$$

7.5.4.3 [+HIGH] stems: with 1st and 2nd personal pronouns as subjects.

(xx)	(1)	(2)	(3)	(4)
	<u>Verb infinitive</u>	<u>Future tense</u>	<u>With object infix and 1st and 2nd personal pronouns as subjects</u>	<u>Category of infix</u>

/kù-rúmúrá/ → [nd-ì-tʃá-rùmùrà] [nd-ì-tʃá-mù-rúmúrá] "I shall
"to wean" wean him" 3rd person sing.

/kù-pérékéḍà/ → [t-ì-tʃá-pèrèkéḍà] [t-ì-tʃá-kù-pérékéḍà] "we
"to accompany" shall accompany you" 2nd pers. sing

/kù-býárúrá/ → [m-ù-tʃá-býarùrà] [m-ù-tʃá-dzì-býárúrá] "you
"to tear" will tear them" N.10

/kù-séʋénzèsa/ → [ù-tʃá-sèʋénzèsa] [ù-tʃá-jì-séʋénzèsa] "you
"to use" will use it" N.9

We have a process identical to that whereby the past tense interacts with the object infix. The high verb stem is lowered to [-HIGH] in the environment of the high tense marker - here [tʃá]: there is tonal dissimilation in column (2). Rule 59 will now be rewritten as 63 (a) and (b) below:

Rule 63

$$(a) \quad \begin{bmatrix} \text{VERB} \\ +H+H.. \end{bmatrix} \rightarrow [-H-H..] / \begin{bmatrix} -\text{HIGH} \\ \text{Noun} \end{bmatrix} + \begin{bmatrix} +\text{HIGH} \\ \{ \text{PAST} \} \\ \{ \text{FUT.} \} \end{bmatrix} + \text{---}$$

But when the low tone object infix is inserted between the high tense markers and the low verb stem, the latter regains its high tone pattern:

$$(b) \quad \emptyset \rightarrow \begin{bmatrix} -\text{HIGH} \\ \text{O.I} \end{bmatrix} / \begin{bmatrix} -\text{HIGH} \\ \text{Noun} \end{bmatrix} + \begin{bmatrix} +\text{HIGH} \\ \{ \text{PAST} \} \\ \{ \text{FUT} \} \end{bmatrix} + \text{---} + \begin{bmatrix} +H+H.. \\ \text{V.STEM} \end{bmatrix} \quad S$$

7.5.4.4 [+HIGH] stems: with 3rd personal pronouns and [-HUMAN] concords as subjects.

(xx1)	(1)	(2)	(3)	(4)
	<u>Verb infinitive</u>	<u>Future tense</u>	<u>With object infix and 3rd person.pron. and [-HUMAN] concords as subjects</u>	<u>Category of infix</u>
/kù-rúmúra/	→	[á-tʃà-rúmúra]	[á-tʃà-mú-rúmúra]	"she'll wean him" 3rd person pron. sing.
/kù-pérékéǵà/	→	[ǵ-á-tʃà-pérékéǵà]	[ǵ-á-tʃà-kú-pérékéǵà]	"they'll accompany you" singular.
/kù-rúmúra/	→	[ǵ-í-tʃà-rúmúra]	[ǵ-í-tʃà-jí-rúmúra]	"they'll wean it" N.9

The derivation of the sentence with the object infix takes place in two stages. First, the tense marker becomes low in the environment [+HIGH] subject to the left and [-HIGH] object infix to the right:

Rule 64

$$(a) \quad \begin{bmatrix} +HIGH \\ \{ PAST \} \\ \{ FUT. \} \end{bmatrix} \rightarrow [-HIGH] / \quad \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \text{---} + \begin{bmatrix} -HIGH \\ \text{O.I.} \end{bmatrix} + \begin{matrix} +H+H.. \\ \text{V.STEM} \end{matrix} \parallel \quad S$$

Secondly, the object infix assimilates to the tone of the initial syllable of the high stem verb:

$$(b) \quad \begin{bmatrix} -HIGH \\ \text{O.I} \end{bmatrix} \rightarrow [+HIGH] / \quad \begin{bmatrix} +HIGH \\ \text{Noun} \end{bmatrix}_{\text{Subj.}} + \begin{bmatrix} -HIGH \\ \{ PAST \} \\ \{ FUT. \} \end{bmatrix} + \text{---} + \begin{matrix} +H+H.. \\ \text{V.STEM} \end{matrix} \parallel \quad S$$

APPENDIX I

The tone patterns of derived nominals (see sections 6.3.1 and 6.3.2).

Although in those sections our main concern is with the influence of the prefixes /N-/ and /ri-/ on the segments at the beginning of verb stems, an important part of those derivations is the assignment of tone to the derived nominals. This is done by rules of the phonological component. An outline of those rules will now be given.

Two classes of nominals were discussed: those belonging to class 9/10, and those belonging to class 5. There are no differences in tone pattern between these classes. The examples used fall into two categories, the regular and the irregular. Thus all the examples marked (i) in 6.3.1 are regular in that the tone pattern of the verb is retained in the nominal derived from it:

<u>Verb</u>	<u>Nominal</u>
/+H+H/	[+H+H]
/-H-H/	[-H-H]
/-H+H/	[-H+H]
/-H-H-H/	[-H-H-H]

The examples marked (ii) are also all regular. Since the nominals are all derived from verb stems, they fall into two tonal classes: the [+HIGH] and the [-HIGH]. Each [+HIGH] verb stem with three syllables yields a nominal with [+HIGH] on the first two syllables, the third

being lowered to [-HIGH]:

<u>Verb</u>		<u>Nominal</u>
/-pópótá/	→	[mḥópótò]
/-páridzá/	→	[mḥáridzò]

When the verb has two [+HIGH] syllables, they retain that pattern in the derived nominal, e.g.:

/-péngá/	→	[mḥéngó]
----------	---	----------

Low stems retain their tone pattern irrespective of the number of syllables involved

/-H-H/	→	[-H-H]
/-H-H-H/	→	[-H-H-H]

The examples marked (iii) and (iv) have some exceptions to these general rules:

(iii)	/-túm ^w á/	→	[ṇḥúm ^w à]
	/-táp ^x á/	→	[ṇḥáp ^x à]
(iv)	/-kóká/	→	[ḥókà]

All [+HIGH] verbs with four syllables, i.e. with the tone pattern [+H+H+H-H], retain this pattern in the derived nominals:

(iv)	/-kúrukúra/	→	[ḥúrukúro]
	/-kóromóra/	→	[ḥóromóri]
(v)	/-wá ^w arí ^r à/	→	[mbá ^w arí ^r o]
	/-wángárádza/	→	[mbángárádzo]

The only other exceptions, i.e. the irregular examples, are found in examples marked (vii):

	/-byóngódzá/	→	[myòng ^w é]
	/-byá ^r urá/	→	[myá ^r u-myá ^r u]
and (viii)	/-kí ^r á/	→	[ṇkí ^r é]

(ix)	/-pèsànà/	—————→	[ɓèsán ^w à]
(x)	/-táng ^w á/	—————→	[dáng ^w è]
	/-tám ^w bá/	—————→	[dám ^w bè]

Rule 1 (iii) finally, the initial syllable of the [-HIGH] verb stem is raised to [+HIGH], i.e. assimilates to the tone of the PAST morpheme which immediately precedes it:

$$[-H-H...]_{Vb} \longrightarrow [+H-H...] / [Noun]_{Subj} + \begin{bmatrix} +HIGH \\ \{PAST\} \end{bmatrix} + \text{---} \Big]_S$$

N.B. this happens irrespective of the tone of the [Noun]_{subj.} in as far as [-HIGH] verb stems are concerned.

(b) [+HIGH] stems

The tone of the [Noun]_{subj.} has to be taken into account where [+HIGH] stems are concerned. Throughout our discussion, the 1st and 2nd personal pronouns functioning as the subject of the verb were treated separately, as having [-HIGH] underlying tone, and as being responsible for tonal processes distinct from those caused by the 3rd personal pronoun and all concords referring to [-HUMAN] nouns, which have [+HIGH] underlying tone.

(i) with 1st and 2nd personal pronouns functioning as subject.

<u>Infinitive</u>	<u>Past tense</u>
/kù-rù'mù'rá/ \longrightarrow	[nd-à-ká-rù'mù'rá]
/kù-tén'gá/ \longrightarrow	[t-à-ká-tén'gá]
/kù-sé'sen'zè'sà/ \longrightarrow	[w-à-ká-sé'sen'zè'sà]
/kù-byà'rù'rá/ \longrightarrow	[m-à-ká-byà'rù'rá]

i.e., after the application of the preliminary Rules 1(i) and (ii) above, the [+HIGH] tones of the verb stems are all lowered:

Rule 2.

$$[+H(+H....)]_{Vb} \longrightarrow [-H(-H....)] / \begin{bmatrix} -HIGH \\ Noun \end{bmatrix}_{Subj} + \begin{bmatrix} +HIGH \\ \{PAST\} \end{bmatrix} + \text{---} \Big]_S$$

(ii) with 3rd personal pronoun and [-HUMAN] concords
functioning as subject.

When, however, [Noun]_{subj.} is the 3rd personal pronoun or a concord referring to [-HUMAN] nouns, i.e. is [+HIGH], the situation is different from that summarised by Rule 2:

[(mùḍzìmaí)]	á-kà-rúmúra]
[(ʋàkúru)]	ʋ-á-kà-byáru]
[(ḡòmbè)]	j-á-kà-túngá]
[(tʃítíma)]	tʃ-á-kà-pútsá]

i.e. in this case, the verb stem retains its underlying [+HIGH] tone pattern. But the {PAST} morpheme dissimilates to that and to the tone of the [Noun]_{Subj.}:

Rule 3.

$$\left[\begin{array}{c} +\text{HIGH} \\ \{ \text{PAST} \} \end{array} \right] \longrightarrow [-\text{HIGH}] \quad / \quad \left[\begin{array}{c} +\text{HIGH} \\ \text{Noun} \end{array} \right]_{\text{Subj.}} + \text{---} + \text{+H(+H...)} \quad \Bigg]_S$$

Rules 2 and 3 show that for [+HIGH] stem verbs, [Noun]_{subj.} bears the same tone in the phonetic representation as the stem of the verb.

Rule 4.

$$[\text{Noun}]_{\text{subj.}} \longrightarrow \left[\begin{array}{c} \alpha\text{HIGH} \end{array} \right] / \text{---} + \left[\begin{array}{c} +\text{HIGH} \\ \text{PAST} \end{array} \right] + \alpha\text{H}(\alpha\text{H...}) \quad \Bigg]_S$$

As with the past tense inflection, high stem verbs have to be treated under two subsections: when the subject Noun is the 1st or 2nd person (singular or plural) pronoun, i.e. [-HIGH], and when the subject Noun is the 3rd personal pronoun (sing. or plural) or any [-HUMAN] concord.

- (i) With the 1st and 2nd personal pronoun functioning as subject.

<u>Infinitive</u>		<u>Future tense</u>
/kù-rù'mù'rá/	→	[nd-ì-tʃá- rù'mù'rá]
/kù-té'ngá/	→	[m-ù-tʃá-té'ngá]
/kù-sé'vénzèsà/	→	[ù-tʃá-sé'vénzèsà]
/kù-býá'rú'rá/	→	[t-ì-tʃá-býá'rú'rá]

The rule operating here is the same as Rule 2 above, dealing with the past tense inflection. Rule 6 will therefore incorporate both inflections:

Rule 6.

$$[+H(+H...)] \rightarrow [-H(-H...)] / \left[\begin{array}{c} [-HIGH] \\ [Noun] \end{array} \right]_{\text{Subj}} + \left[\begin{array}{c} [+HIGH] \\ \{FUT\} \\ \{PAST\} \end{array} \right] + \left[\begin{array}{c} \text{---} \\ S \end{array} \right]$$

- (ii) With the 3rd personal pronoun and [-HUMAN] concords functioning as subject.

<u>Infinitive</u>		<u>Future tense</u>
/kù-rù'mù'rá/	→	[á-tʃá-rù'mù'rá]
/kù-té'ngá/	→	[v-á-tʃá-té'ngá]
/kù-sé'vénzèsà/	→	[r-ì-tʃá-sé'vénzèsà]
/kù-býá'rú'rá/	→	[dʒ-ì-tʃá-býá'rú'rá]

The rule here is the same as Rule 3 above. So Rule 7 will incorporate both inflections:

Rule 7.

$$\left[\begin{array}{c} [+HIGH] \\ \{FUT\} \\ \{PAST\} \end{array} \right] \rightarrow [-HIGH] / \left[\begin{array}{c} [+HIGH] \\ [Noun] \end{array} \right]_{\text{Subj.}} + \text{---} + H(+H...) \left[\begin{array}{c} \text{---} \\ S \end{array} \right]$$

i.e. the tense marker dissimilates from the surrounding [+HIGH] syllables.

BIBLIOGRAPHY

- Abercrombie, D. et al. (eds)(1964) In Honour of Daniel Jones. Longmans, Green & Co. Ltd.
- _____ (1967) Elements of General Phonetics, Edinburgh University Press.
- Allen, J.P.B. & Van Buren, P. (eds)(1971) Chomsky: Selected Readings, O.U.P.
- Anderson, S.R.(1974) The Organization of Phonology, Academic Press.
- Arnold, G.F. (1966) "Concerning the theory of Plosives", in Jones & Laver (eds) Phonetics in Linguistics: a book of Readings (1973).
- Beach, D.M. (1924) "The Science of Tonetics and its Application to Bantu Languages" Bantu Studies Vol.11, No.2.
- _____ (1938) The Phonetics of the Hottentot Language, Heffer & Sons Ltd. Cambridge.
- Bennett, D.C. (1968) "Spectral Form and Duration as Cues in the Recognition of English and German Vowels" Language and Speech, 11.
- Bloch, B. (1948) "A set of postulates for phonemic analysis" Language, 24, pp.3-46.
- _____ and Trager, G.L. (1941) "The Syllabic Phonemes of English". Language, Vol.17.
- _____ (1942) An Outline of Linguistic Analysis, Baltimore.
- Bloomfield, L. (1933) Language, 8th ed. Allen & Unwin, London.
- Bolinger, D. (1968) Aspects of Language, Harcourt, Brace & World Inc.
- Botha, R.P. (1971) Methodological Aspects of Transformational Generative Phonology. Mouton.
- _____ (1973) The Justification of Linguistic Hypotheses: a study of nondemonstrative inference in transformational grammar. Mouton.
- Brosnahan, L.F. & Malmberg, B.(1970) Introduction to Phonetics, Heffer & Sons Ltd. Cambridge.
- Carnochan, J. (1960) "Vowel Harmony in Igbo" African Language Studies, Vol.1, pp.155-63.
- _____ (1962) "Pitch, Tone and Intonation in Igbo" Proceedings of the 4th International Congress of Phonetics Sciences, Helsinki 1961, Mouton & Co.

- Carnochan, J. (1964) "Pitch, Tone and Intonation in Yoruba"
In Abercrombie et al (eds) 1964.
- _____ (1966) "Word Classes in Igbo" Lingua, 17, pp.1-23
- Carter, Hazel J. (1956) "Stabilization in the Manyika dialect
of the Shona Group" Africa, XXVI, 4, pp.398-405
- _____ (1975) "The Abstract in Bantu Languages"
Paper read at International Conf. on Lings.
at Univ. of Rhodesia.
- Catford, J.C. (1939) "On the classification of Stop Consonants"
In Jones & Laver (eds) 1973.
- _____ (1964) "Phonation Types: the classification of
laryngeal components of speech production"
In Abercrombie et al (eds) 1964.
- _____ (1970) "The Articulatory Possibilities of Man"
In Malmberg, B (ed) Manual of Phonetics,
North-Holland.
- _____ (1976) Fundamental Problems in Phonetics
Edinburgh University Press.
- Chomsky, N. (1957) Syntactic Structures. The Hague, Mouton.
- _____ (1965) Aspects of the Theory of Syntax, M.I.T.
Press
- _____ (1966) "The current scene in Linguistics: present
directions" In Reibel & Schane (eds) Modern
Studies English 1969.
- _____ (1976) Reflections on Language, Temple Smith.
- Chomsky, N. and Halle, M. (1965) "Some controversial Issues in
Phonological Theory" Journal of Linguistics 1,
pp.97-138.
- _____ (1968) The Sound Pattern of English
Harper & Row, New York.
- Cole, D.T. (1971) "The History of African Linguistics to
1945" In T.A. Sebeok (ed) Current Trends in
Linguistics, Vol.7
- Delattre, P. (1968) "From Acoustic Cues to Distinctive
Features" Phonetica, 18, pp.198-230.
- _____ et al (1952) "An Experimental Study of the
Acoustic Determinants of Vowel Color"
Word, 8, 195.
- Dembetembe, N.C. (1969) Verbal Constructions in Korekore
Dialect. University of London M.Phil. thesis.
Unpublished.

- Dembetembe, N.C. (1974) "Three Kore Kore Dialects" African Language Studies, XV, 143-172
- _____ (1976) The Syntax of Sentential Complements in Shona. Univ. of London Ph.D. thesis, Unpublished.
- Denes, P.B., & Pinson, E.N. (1963) The Speech Chain: the physics and biology of spoken language. Bell Telephone Laboratories 8th Printing, 1972.
- Derwing, B.L. (1973) Transformational Grammar as a theory of Language Acquisition, Cambridge Univ. Press.
- Doke, C.M. (1931) Report on the Unification of the Shona Dialects, Austin & Son.
- _____ (1931) A Comparative Study in Shona Phonetics University of the Witwatersrand Press
- Fant, G. (1960) Acoustic theory of speech production, The Hague.
- _____ (1970) "Analysis and Synthesis of Speech Processes" In B. Malmberg (ed) 1970.
- _____ (1973) Speech Sounds and Features, M.I.T. Press.
- Firth, J.R. (1957) Papers in Linguistics, 1934-51, O.U.P.
- Fischer-Jørgensen, E. (1975) Trends in Phonological Theory: a historical introduction. Akademisk Forlag. Copenhagen.
- Fivaz, D. (1970) Shona Morphophonemics and Morphosyntax. Univ. of the Witwatersrand Press.
- Fortune, G. (1955) An Analytical Grammar of Shona, Longmans
- _____ (1968) Shona Grammatical Constructions, Vols. I & II (forthcoming).
- _____ (1970) "The reference of Primary and Secondary Noun Prefixes in Zezuru" African Studies, 29, pp.81-110.
- Fromkin, V.A. (1976) "A Note on Tone and the Abstractness Condition" Studies in African Linguistics, Supplement 6, Nov. 1976, pp.47-62, UCLA.
- Fry, D.B. (1960) "Linguistic Theory and Experimental Research" Transactions of the Philological Society, pp.13-39.
- _____ (ed)(1977) Acoustic Phonetics: a course of basic readings, Cambridge University Press.

- Fudge, E.C. (1967) "The Nature of Phonological Primes"
Journal of Linguistics, 3, pp.1-36.
- _____ (ed)(1973) Phonology, Penguin Books.
- Fujimura, O. (1962) "Analysis of Nasal Consonants" Journal of the Acoustical Society, 34, pp.1865-75
- Gimson, A.C. (1962) An Introduction to the Pronunciation of English. Edward Arnold, Reprinted 1972.
- _____ (1945-49) "Implications of the phonemic/chronemic grouping of English Vowels"
Acta Linguistica Vol.V, Fasc.2: pp.94-100
- _____ (1956) "The Linguistic Relevance of Stress in English" In Jones & Laver (eds) pp.94-102.
- Gleason, H.A.Jr. (1956) Review of Southern Bantu Languages by C.M. Doke, Language, 32.
- _____ Review of An Analytical Grammar of Shona, by G. Fortune. Language, 32.
- Halle, M., Hughes, G.W. & Radley, J.P. (1957) "Acoustic Properties of Stop Consonants" Journal of the Acoustical Society of America (J.A.S.A) 29, 107.
- Hannan, M., S.J. (1974) Standard Shona Dictionary. Rhodesia Literature Bureau.
- Harris, K.S. (1954) "Cues for the Identification of the Fricatives of American English" J.A.S.A. 26, 252.
- _____ (1958) "Cues for the Discrimination of the Fricatives of American English in Spoken Syllables" Language & Speech, 1.1.
- Harris, Z. (1951) Structural Linguistics, Univ. of Chicago Press, 8th Impression, 1969.
- Heffner, R-M.S. (1950) General Phonetics, Univ. of Wisconsin Press.
- Hempel, C. (1966) Philosophy of Natural Science, Englewood Cliffs, N.J.: Prentice-Hall.
- Henderson, E.J.A. (1965) The Domain of Phonetics. University of London Inaugural Lecture, O.U.P.
- _____ (ed)(1971) The Indispensable Foundation: a selection from the writings of Henry Sweet O.U.P.

- Hjelmslev, L. (1943) Prolegomena to a theory of Language.
Translated by F.J. Whitfield, Univ. of
Wisconsin Press, 1961.
- Hockett, C.F. (1942) "A system of descriptive phonology"
Language 18, Reprinted in V.B. Makkai (ed)
pp.99-112
- Hooper, J.B. (1973) Aspects of Natural Generative Phonology
Univ. of California, Los Angeles, Ph.D.
thesis.
- Householder, F.W. (1965) "On some recent claims in phonolog-
ical theory" Journal of Linguistics 1, pp.13-34.
- Hyman, L.M. (1973b) "The Role of Consonant Types in natural
tonal assimilations" In L.M. Hyman (ed)
Consonant Types and Tone, Southern California
Occasional Papers in Linguistics, No.1
- _____ (1976) Phonology: Theory and Analysis. Holt, Rinehart &
Winston.
- Jakobson, R. et al (1951) Preliminaries to Speech Analysis.
The Distinctive Features and their correlates.
M.I.T. Press, 2nd edn. 1963.
- Jones, D. (1918) An Outline of English Phonetics Heffer &
Sons Ltd. 9th Edn, 1969.
- _____ (1931) "The 'word' as a phonetic entity" Le Maître
Phonétique, No.36:60-65.
- _____ (1944a) "Chronemes and Tonemes, a contribution to
the study of the theory of phonemes" Acta
Linguistica, Vol.1V, Fasc1:1-10
- _____ (1944b) "Some Thoughts on the Phoneme" Transactions
of the Philological Society, 119-135.
- _____ (1948) "The London School of Phonetics" Reprinted
in Jones & Laver (eds) 1973.
- _____ (1950) The Phoneme: its nature and use. 3rd Edn.
1967, Heffer & Sons
- _____ (1957) "The history and meaning of the term
'phoneme'" Appendix to D. Jones, 3rd ed (1967)
- Jones, W.E. and Laver, J. (eds) (1973) Phonetics in Linguistics:
a book of Readings, Longman
- Kaplan, A. (1964) The Conduct of Inquiry: Methodology for
Behavioural Science, San Francisco, Chandler
Publishing Co.
- Koestler, A. (1976) "The Vision that links the poet, the
painter and the scientist" Opening address
to PEN Conference, The Times Newspaper,
August 25th.

- Kuhn, T.S. (1962) The Structure of Scientific Revolutions
Univ. of Chicago Press. Foundations of
the Unity of Science.
- Ladefoged, P. (1960) "The value of Phonetic Statements"
In Jones & Laver (eds), 1973, pp.218-228.
- _____ (1962) Elements of Acoustic Phonetics,
Univ. of Chicago Press, 9th Impression, 1973.
- _____ (1964a) A Phonetic Study of West African
Languages, Cambridge Univ. Press.
- _____ (1967) Three Areas of Experimental Phonetics,
O.U.P.
- _____ (1971) Preliminaries to Linguistic Phonetics,
Univ. of Chicago Press.
- _____ (1975) A Course in Phonetics, Harcourt, Brace
& Jovanovich.
- Laver, J. (1974) "Labels for Voices" Journal of the Inter-
national Phonetic Association, pp.62-75.
- Lyons, J. (1962) "Phonemic and Non-phonemic phonology: some
typological reflections" In Jones & Laver
(eds) 1973, pp.229-239.
- _____ (1968) An Introduction to theoretical linguistics.
London: Cambridge Univ. Press.
- _____ (1972) Structural Semantics: an analysis of part
of the vocabulary of Plato. Basil Blackwell
Oxford.
- Maddieson, I. (ed)(1970) Tone in Generative Phonology:
Research Notes Vol.3, Parts 2 and 3.
Dept. of Linguistics & Nigerian Languages,
Univ. of Ibadan.
- _____ (1970) "The Inventory of Features required
for handling Tone in Transformational
Phonology:" In Maddieson (ed) 1970.
- Magee, B. (1973) Popper, Fontana.
- Makkai, V.B. (1970) Phonological Theory: evolution and
current practice. Rinehart and Winston.
- Malmberg, B. (ed)(1970) Manual of Phonetics, North-Holland.
- Mandelbaum, D.G. (ed)(1966) Edward Sapir: Culture, Language
and Personality: Selected Readings, Univ. of
California Press.
- Martinet, A. (1960) Elements of General Linguistics, Translated
by E. Palmer, London, Faber & Faber, 1964.

- Matthews, P.M. (1974) Morphology: An Introduction to the Science of Word Formation, Cambridge Univ. Press.
- Meyers, L. (1976) Aspects of Hausa Tone, UCLA Ph.D.thesis, unpublished.
- Mitchell, T.F. (1975) Principles of Firthian Linguistics, Longman.
- Mkanganwi, K.G. (1972) "Relationship of Coastal Ndau to the Shona Dialects of the Interior" African Studies 31,2,pp.111-137.
- _____ (1973) An Outline of the morphology of substantives in Ndau, with a preliminary note on Ndau phonology, Univ. of London M.Phil. thesis.
- _____ (1975) "Towards a description of Town Shona" Paper presented at the International Conference on Linguistics at the Univ. of Rhodesia.
- Ngara, E.A. (1973) The Significance of time and motion in the poetry of T.S. Eliot, with special reference to the teaching of T.S. Eliot in Rhodesia. Univ. of London M.Phil.thesis.
- _____ (1974) "A Redefinition of the Role of the English Language in African Universities" Bulletin of the Association of African Universities, Vol.1, No.2.
- _____ (1975) "The Role of Applied Linguistics in Developing Africa" Paper read at the International Conference on Linguistics at the University of Rhodesia.
- O'Connor, J.D. (1973) Phonetics, A Pelican original.
- _____ et al (1957) "Acoustic Cues for the Perception of initial /w,j,r,l/ in English" Word, 13 pp.24-43.
- Palmer, F.R. (ed) (1970) Prosodic Analysis, Oxford University Press.
- Pike, K.L. (1944) Phonetics: A critical analysis of phonetic theory and a technic for the practical description of sounds. Ann Arbor: University of Michigan Press.
- _____ (1947a) Phonemics: a technic for reducing languages to writing. Ann Arbor, U.M.P.
- _____ (1947b) "Grammatical prerequisites to phonemic analysis", Word 3

- Pike, K.L. (1948) Tone Languages, Ann Arbor, U.M.P.
- _____ (1967) Language in Relation to a Unified theory of Human Behaviour, Mouton and Co. The Hague
- Pongweni, A.J. (1975) The Language of Advertising and its socio-cultural Implications for the Consumer, Univ. of Leeds, M.A. Dissertation.
- Popper, K. (1959) The Logic of Scientific Discovery. Hutchinson, 3rd ed. 1972.
- _____ (1963) Conjectures and refutations: the growth of scientific knowledge, 5th ed. Routledge & Kegan Paul, 1974.
- _____ (1972) Objective Knowledge: an evolutionary approach, Oxford, At the Clarendon Press.
- Postal, P. (1968) Aspects of Phonological Theory, Harper & Row, New York.
- The Principles of the International Phonetic Association, 1949, Reproduced, 1972.
- Robey, D. (ed) (1972) Structuralism: an introduction. Clarendon Press.
- Robins, R.H. (1964) General Linguistics, an introductory survey. Longmans, 2nd edn. 1971.
- _____ (1967) A Short History of Linguistics, Longmans.
- Russell, B. (1946) History of Western Philosophy. Unwin University Books, 8th Impression, 1975.
- Ruwet, N. (1973) An Introduction to Generative Grammar. Translated by N.S.H. Smith. Amsterdam. North-Holland.
- Sapir, E. (1921) Language: an introduction to the study of speech. Rupert Hart-Davis, London, 1970.
- _____ (1925) "Sound Patterns in Language" in Makkai (ed) 1972, pp.12-21.
- _____ (1929) "The Status of Linguistics as a Science" In Mandelbaum (ed) 1966, pp.65-77.
- _____ (1933) "The psychological reality of the phoneme" In Makkai (ed) 1972, pp.22-31.
- Saussure, F. de (1917) Course in General Linguistics. Translated by Wade Baskin 1966, McGraw-Hill Paperbacks.
- Schane, S.A. (1973) Generative Phonology, Prentice-Hall, Inc.

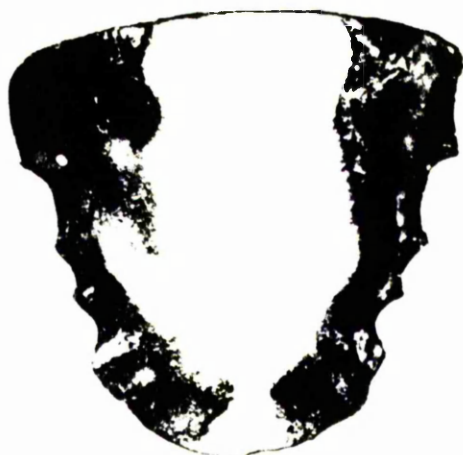
- Sebeok, T.A. (ed)(1971) Current Trends in Linguistics, Vol.7
Mouton.
- Sharp, A.E. (1954) "A tonal analysis of the disyllabic
noun in the Machame dialect of Chaga"
Reprinted in Jones and Laver (eds) 1973,
pp.305-319.
- Spaandonck, Marcel van (1967) Morfotonologische analyse in
Bantutalen. With a summary in English
Tongeren.
- Stevick, E.W. (1960) "The Implosive-Explosive Contrast in
Manyika" African Studies, 19,2, pp.88-95.
- _____ (1964) "Two Bantu Consonant Systems" Language
40:1.
- Stevens, P. (1960) "Spectra of Fricative Noise in Human
Speech" Language and Speech, 3, pp.32-49.
- Swadesh, M. (1934) "The Phonemic Principle" In Makkai (ed)
1972, pp.32-39.
- _____ (1935) "Twaddell on Defining the 'Phoneme'" In
Makkai (ed) 1972, pp.41-44.
- _____ (1937) "The Phonemic interpretation of long
consonants" in Makkai (ed) 1972, pp.51-57.
- Sweet, H. (1877) A Handbook of Phonetics)
_____ (1891) A New English Grammar) Excerpts in
_____ (1911) Phonetics, in Encyclopaedia Britannica; 11th
edn. Vol.21, pp.1910-1911.) Henderson (ed)
1971.
- Trager, G.L. (1950) Review of Phonemics, by K.L. Pike,
Language, 26, 152.
- Trubetzkoy, N.S. (1939) Principles of Phonology. Translated
from German by C.A. Baltaxe, California Univ.
Press, 1969.
- Ward, I.C. (1929) The Phonetics of English, Heffer & Sons,
5th ed. 1972.
- _____ and Westermann, D. (1933) Practical Phonetics for
Students of African Languages, Oxford
Univ. Press.
- Wells, J.C. and Gimson, A.C. (eds)(1974) Journal of the
International Phonetic Association, Vol.4
No.2, December.
- Williamson, K. (1970) "The Generative Treatment of Downstep"
in Maddieson (ed) 1970.



(i)a šimbá
[strength]



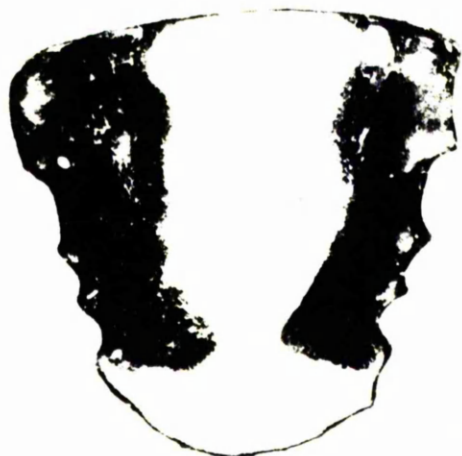
(i)b. -pésá
[pass] (v.t.)



(ii)a. zimbá
[mansion]



(ii)b. -pàza
[destroy]



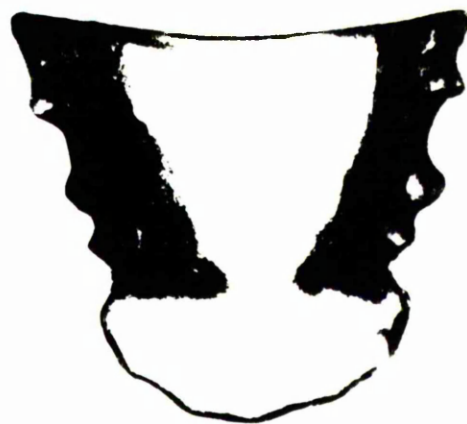
(iii)a. šimbó
[knobkerry]



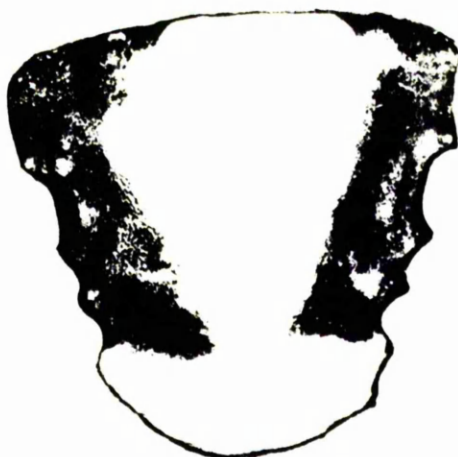
(iii)b. -šavà
[dry or wilt]



(iv)a. ʒipó
[gifts]



(iv)b. ʒámáǵà
[what you have stolen]



(v)a. ʃúmǵà
[lion]



(v)b. -ʃámǵà
[wash]



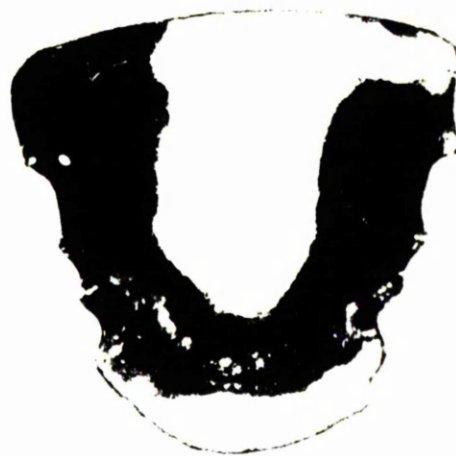
(vi)a. ʒèvé
[ear]



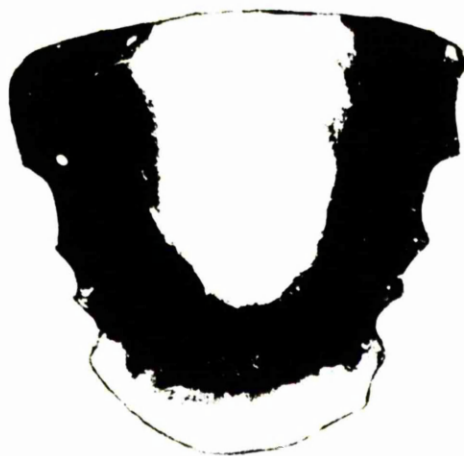
(vi)b. -ʒámǵá
[cry hysterically]



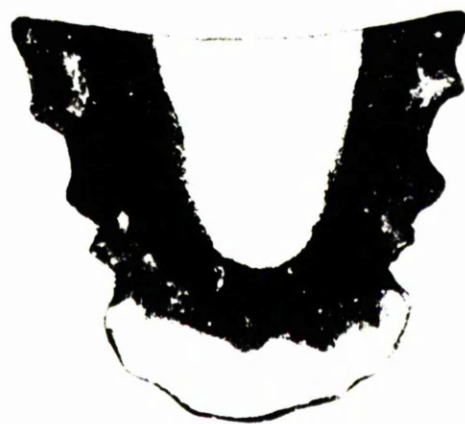
(vii)a. ɛ́iwá
[be feared]



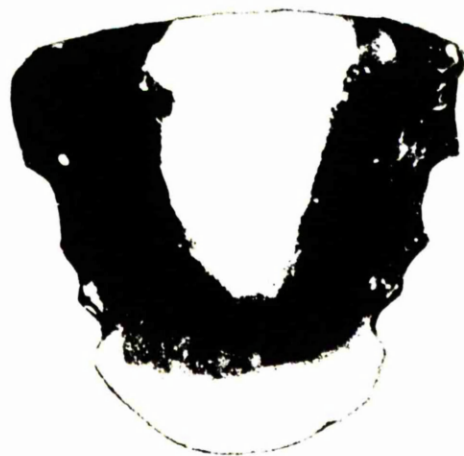
(viii)a. ɣ́iwá
[be eaten]



(ix)a. -témá
[cut or hit with object]



(ix)b. -támá
[move house]



(x)a. díbà
[dip-tank]



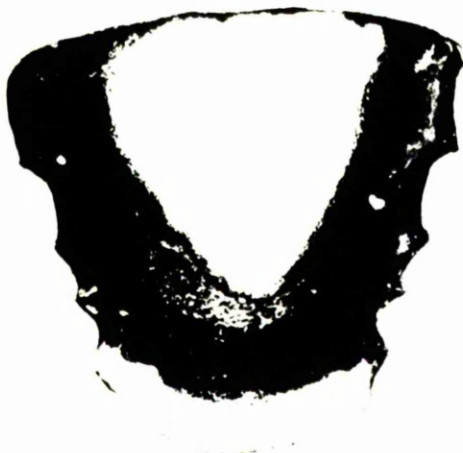
(x)b. dàdà
[duck]



(xi)a. dímà
[darkness of night]



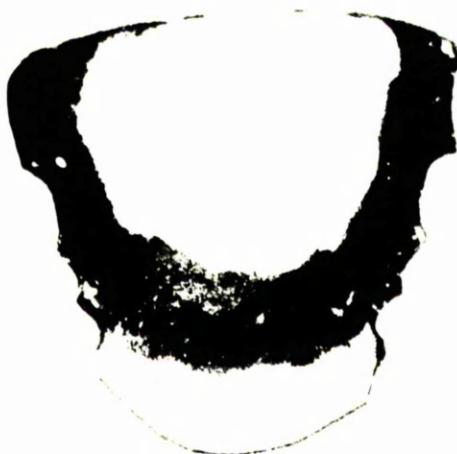
(xi)b. dambá
["fruit"]



(xii)a. ndipe
[give me...]



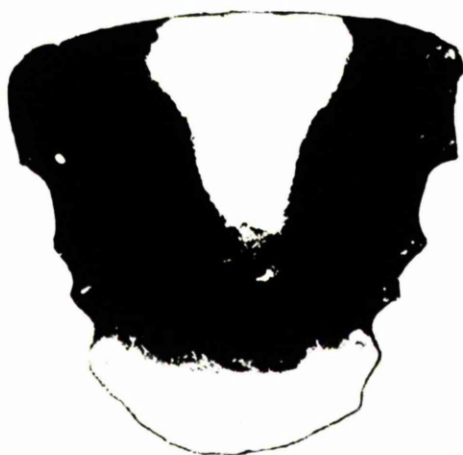
(xii)b. mǎndá
[animal fat]



(xiii)a. -póná
[survive or give birth]



(xiii)b. -páná
[give each other]



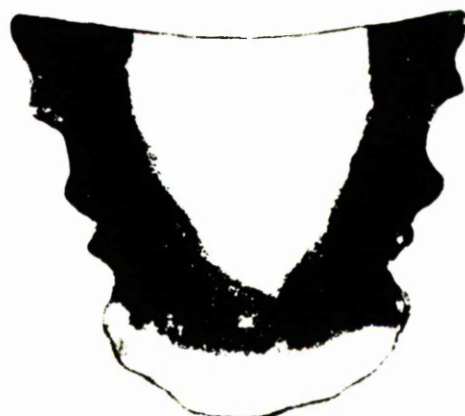
(xiv)a. nìmò
[bambara groundnut]



(xiv)b. pàmà
[meat]



(xv)a. nẓimbè
[sugar cane]



(xv)b. -wànzà
[increase in quantity]



(xvi)a. nẓimbò
[place (n. cl.9)]



(xvi)b. nẓè
[ideo. of carrying on
shoulder]



(xvii)a. n̄kó
[greediness]



(xvii)b. n̄àngà
[grain store]



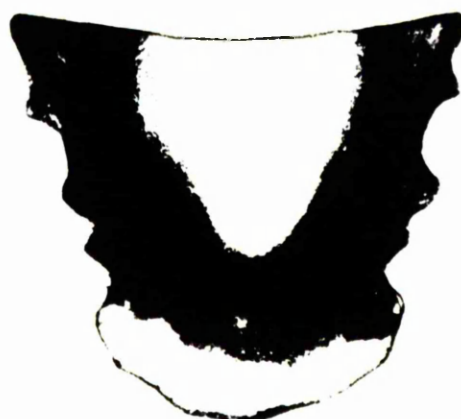
(xviii)a. tsimè
[well (cl.5)]



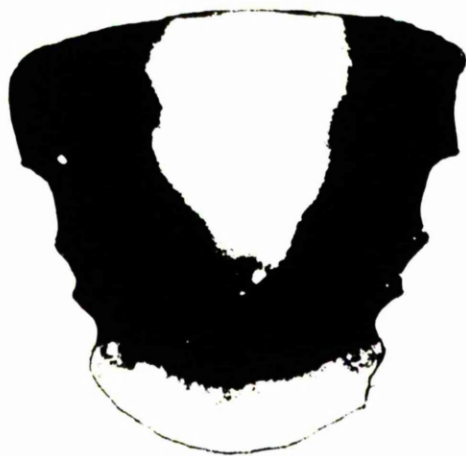
(xviii)b. tsàmbà
[letter]



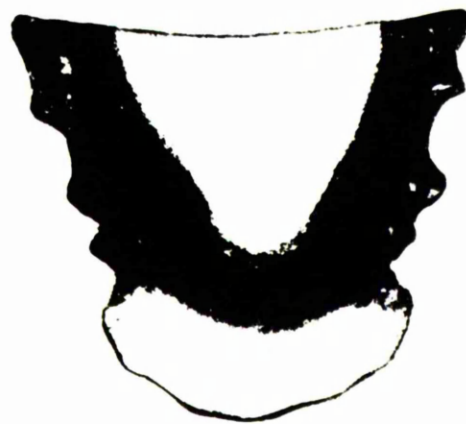
(xix)a. d̄zimba
[houses]



(xix)b. -d̄zàma
[be profound]



(xx)a. `itṣá
[burn!]



(xx)b. mətṣá
[new ones (cl. 6.)]



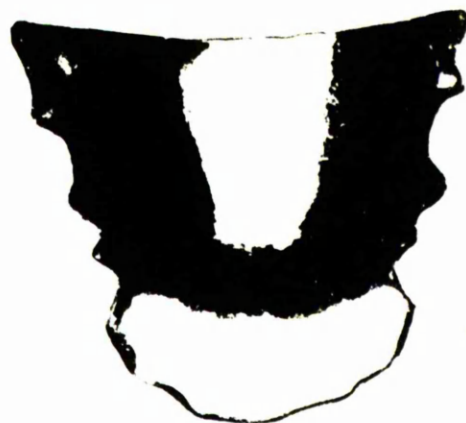
(xxi)a. `iḏṣá
[stumble!]



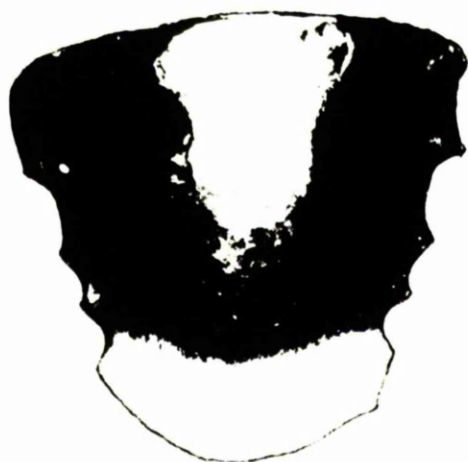
(xxi)b. mədṣá
[huge thighs]



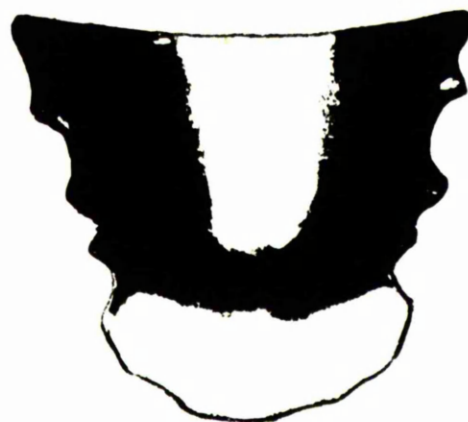
(xxii)a. tṣipó
[gift]



(xxii)b. tṣápá
[careless person]



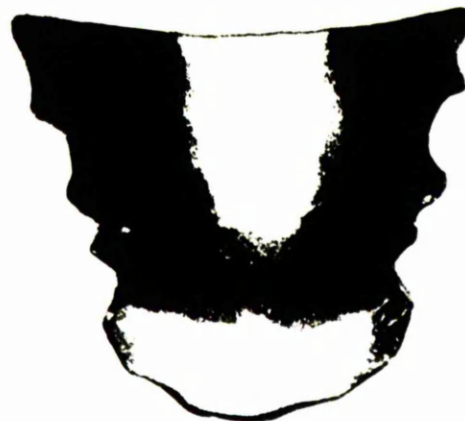
(xxiii)a. m̀dʒiŋǎ
[at the foot of a hill]



(xxiii)b. dʒáǎǎ
[light brown (adj.cl.5)]



(xxiv)a. ndʒivá
[dove]



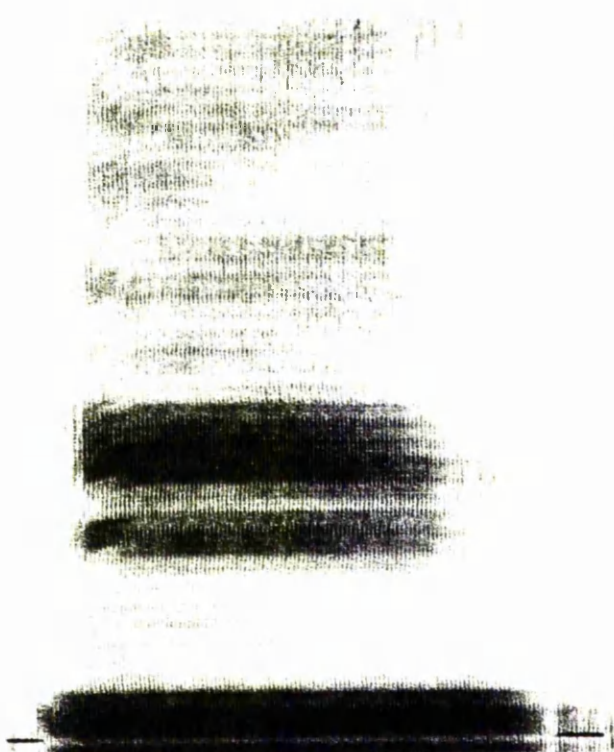
(xxiv)b. m̀andʒǎ
[applause!]



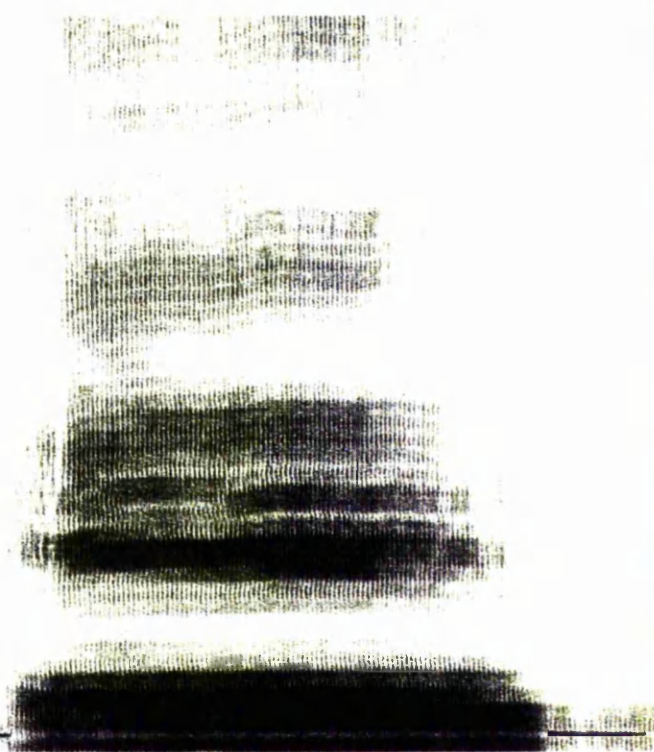
(xxv)a. -rimà
[farm or plough (v.t.&i.)]



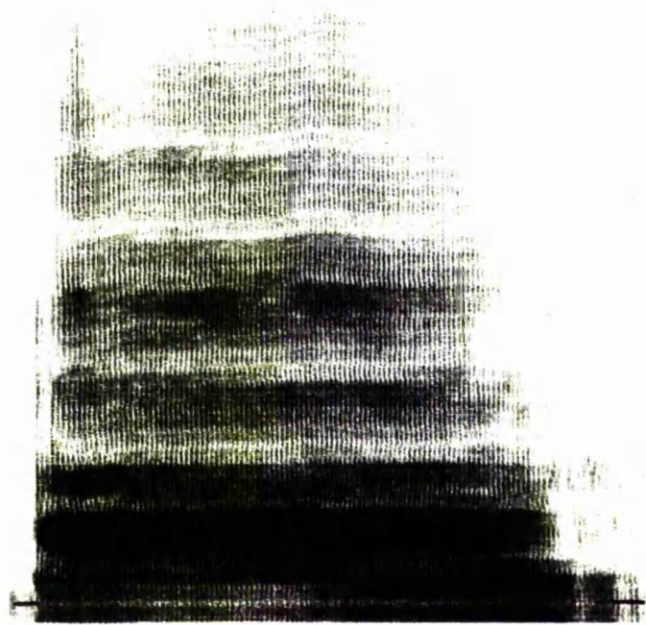
(xxv)b. m̀aràrà
[dirt or refuse]



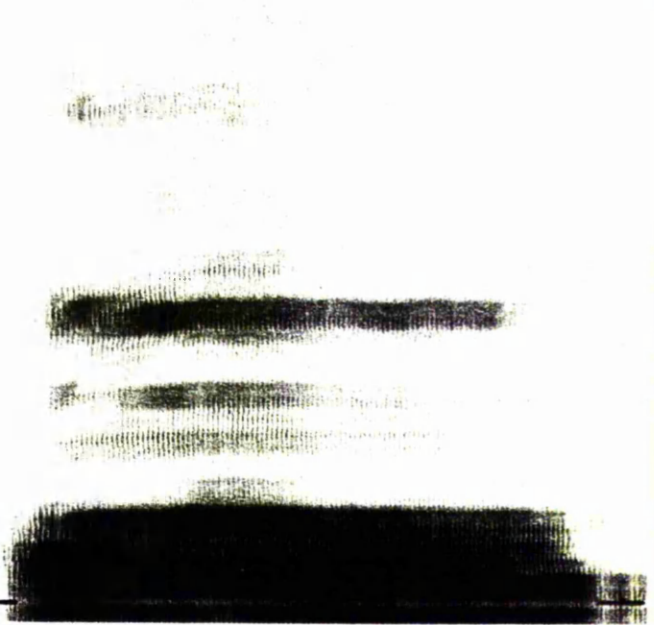
1. [i]



2. [e]



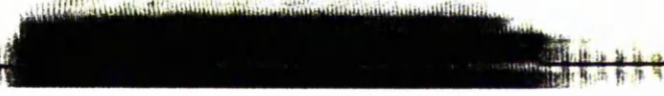
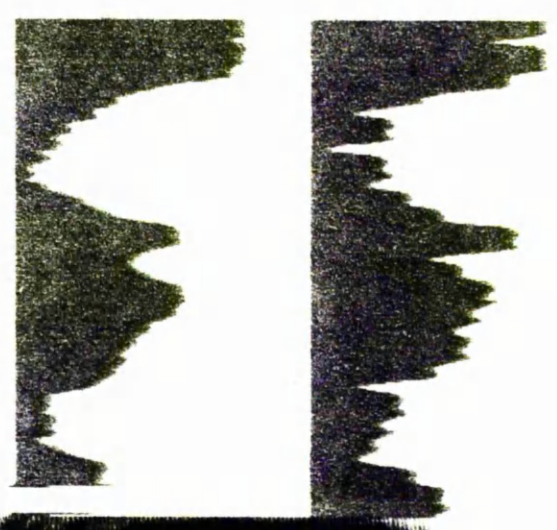
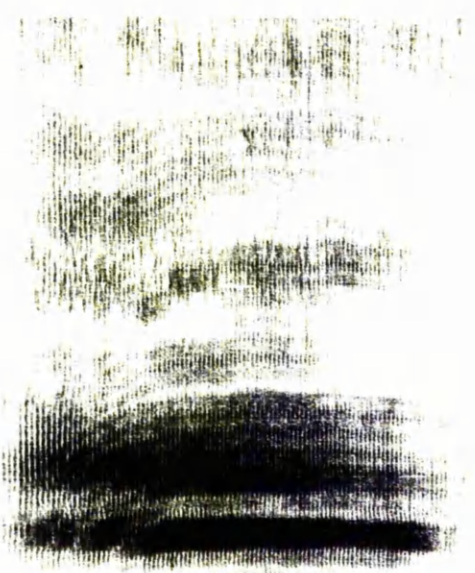
3. [a]



4. [o]



5. [u]



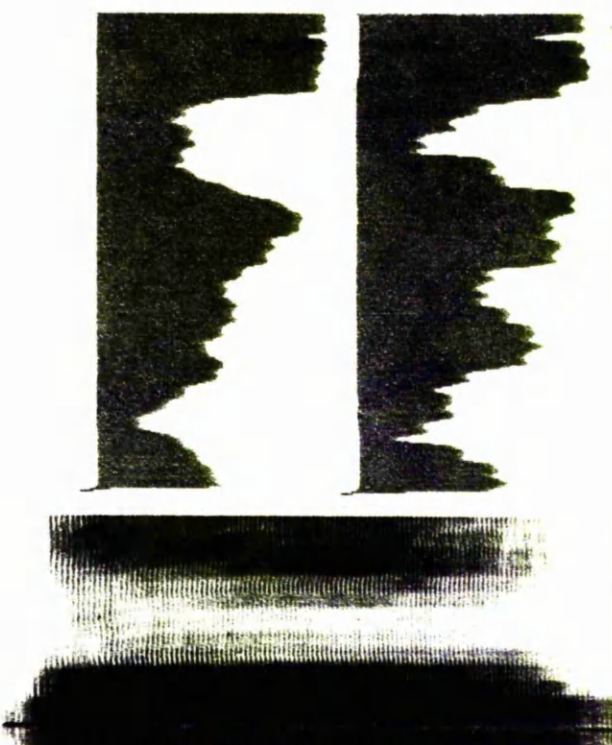
6. h 1

6. h

1 Section



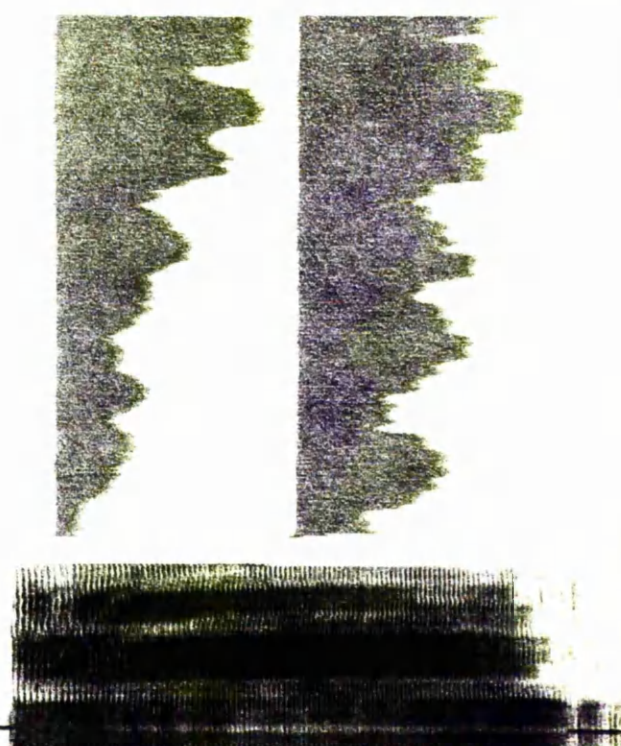
7. \bar{h} \bar{e}



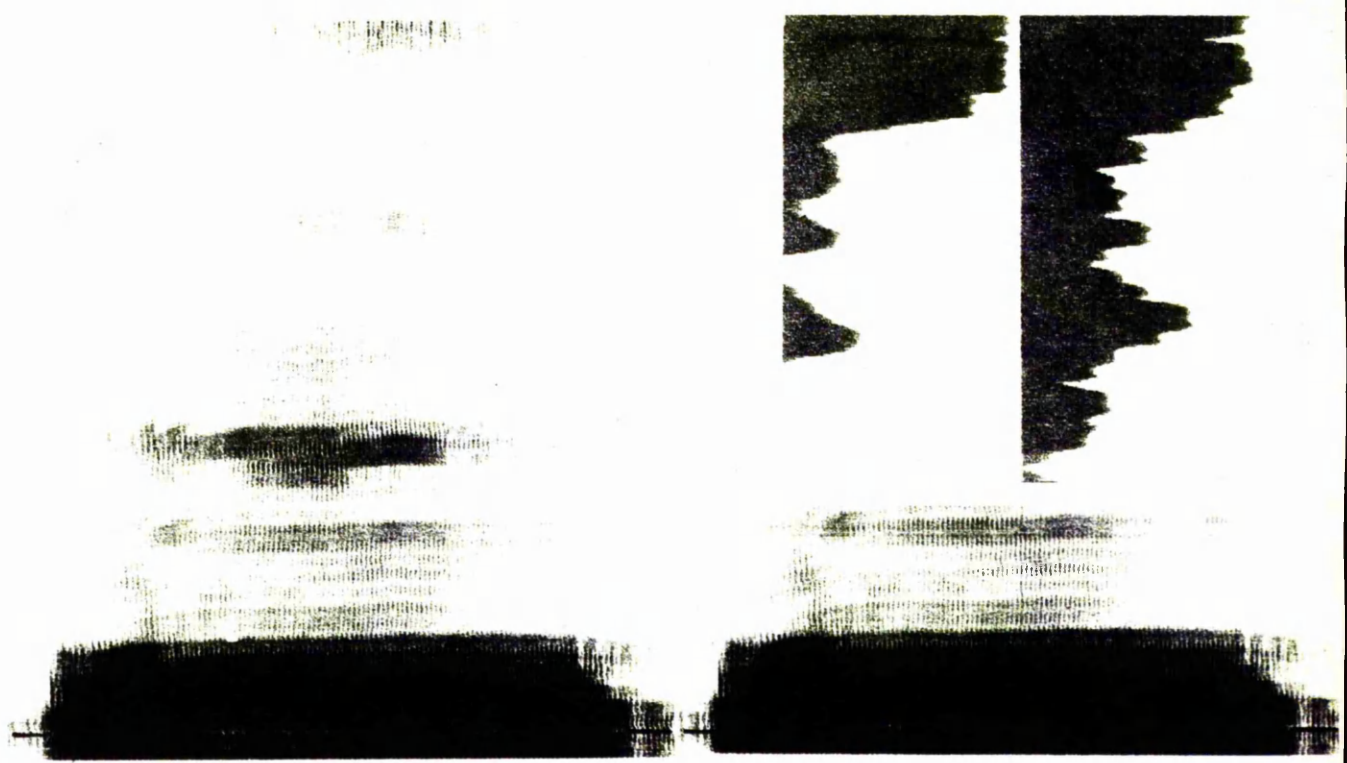
7. \bar{h} \bar{e} Section



8. \bar{h} \bar{a}

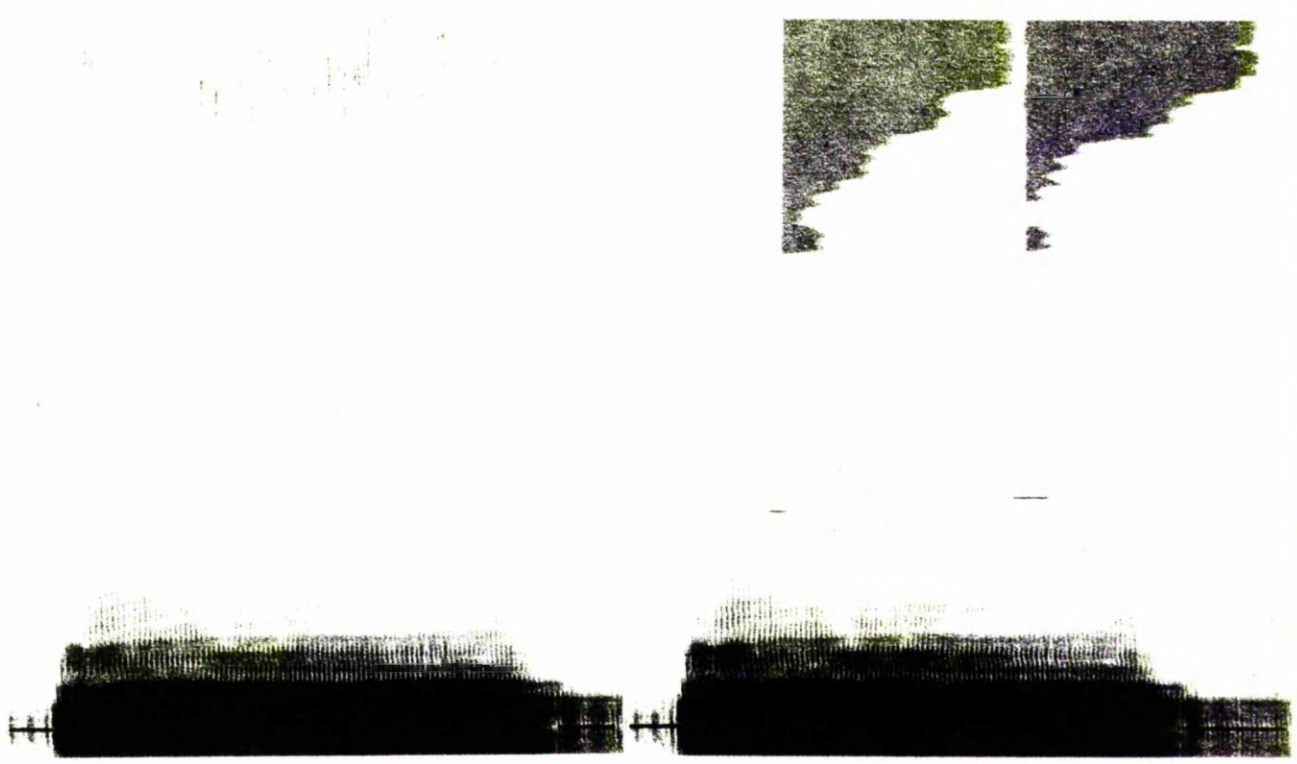


8. \bar{h} \bar{a} Section



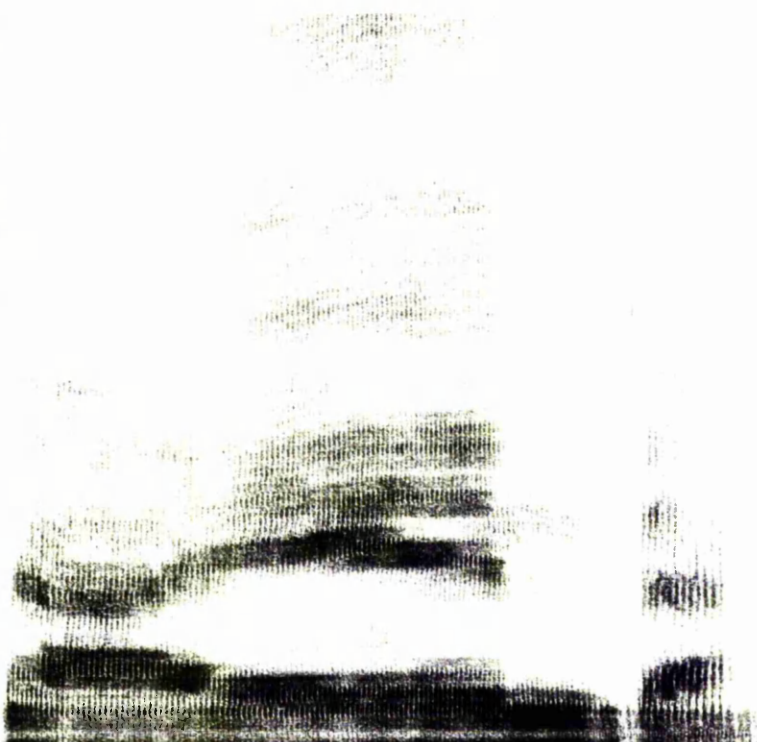
9. \hat{h} \circ

9. \hat{h} \circ Section




10. \hat{h} u

10. \hat{h} u Section



11. t à é n d à



12. t à í m b á



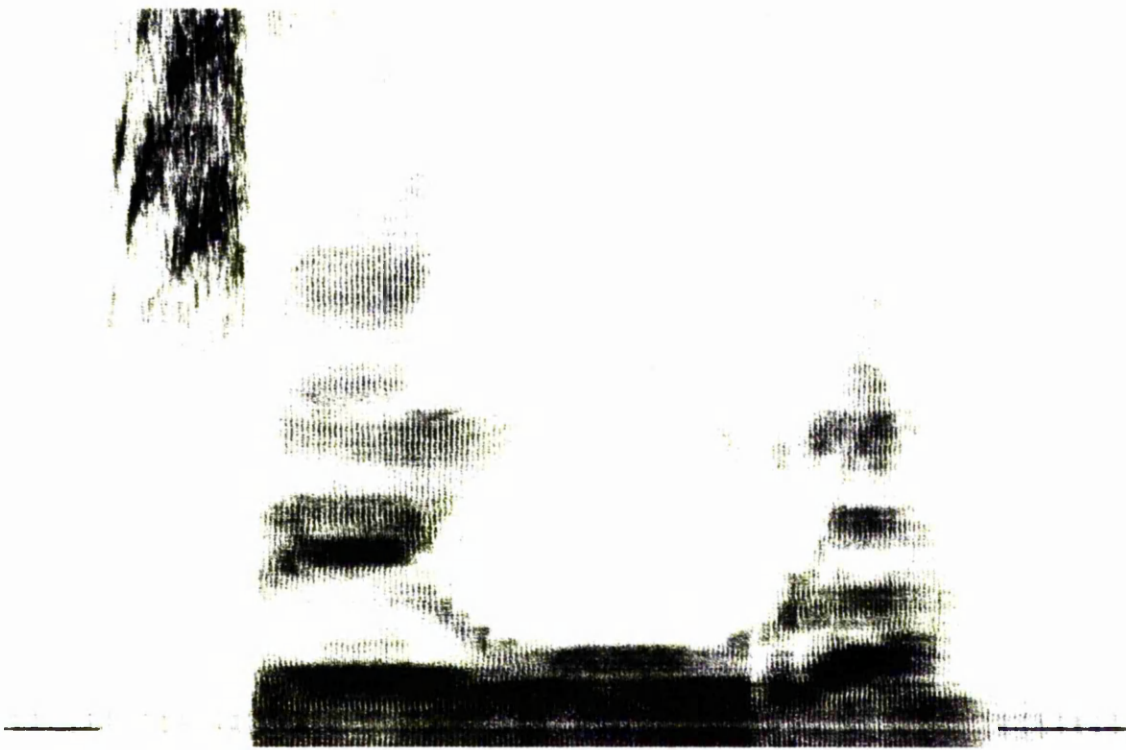
13. t à ó n á



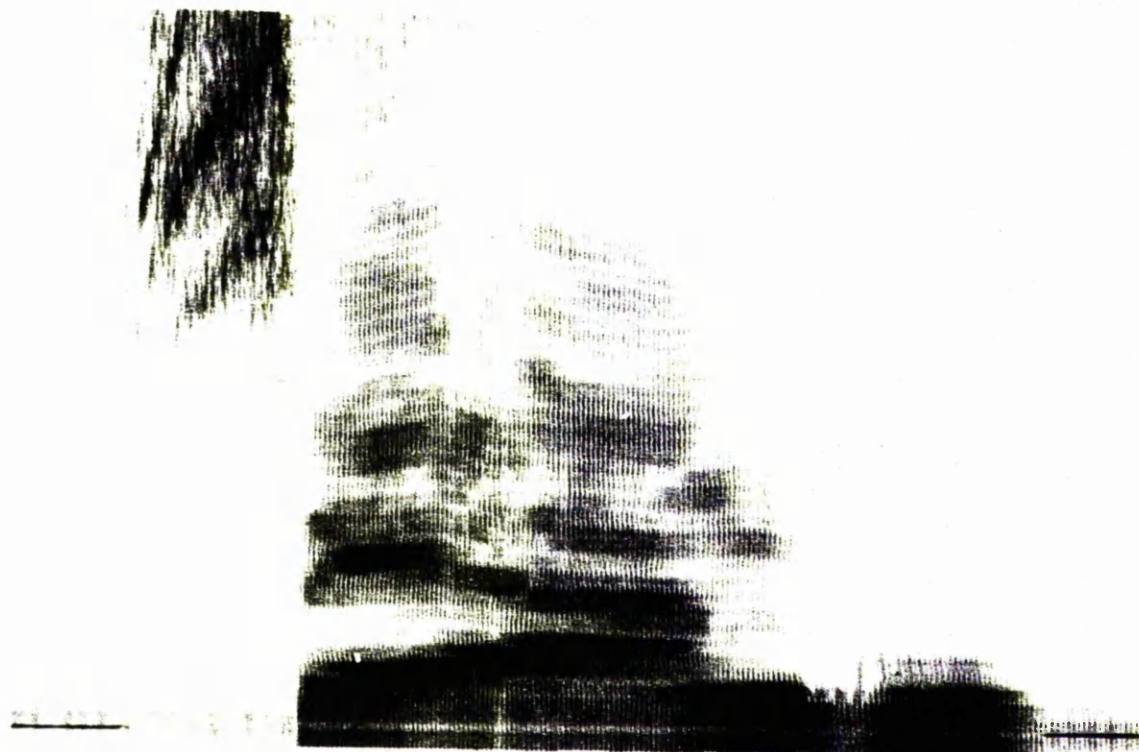
14. t à ú r á



15. s é ì



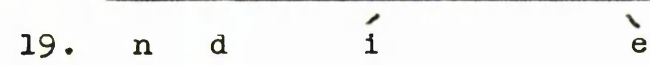
16. s è ú r á



17. s é à n g ú



18. n d í à n í



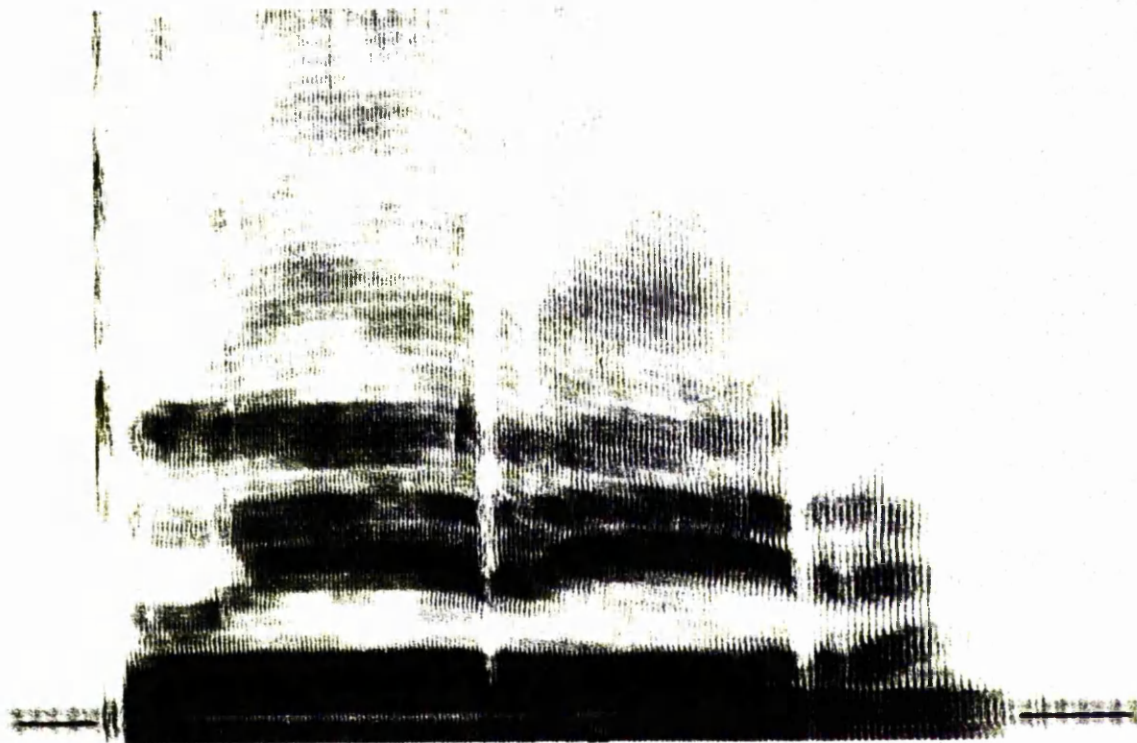
19. n d í è



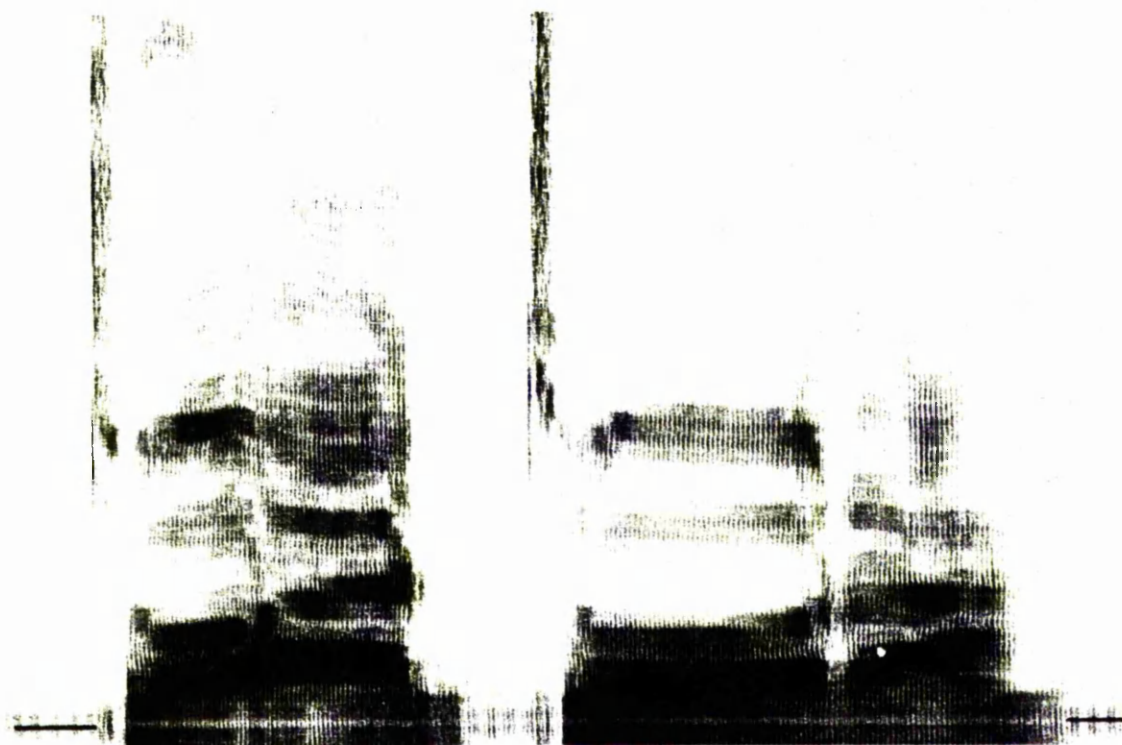
20. n d í ó n è

21. n d í ú m b è

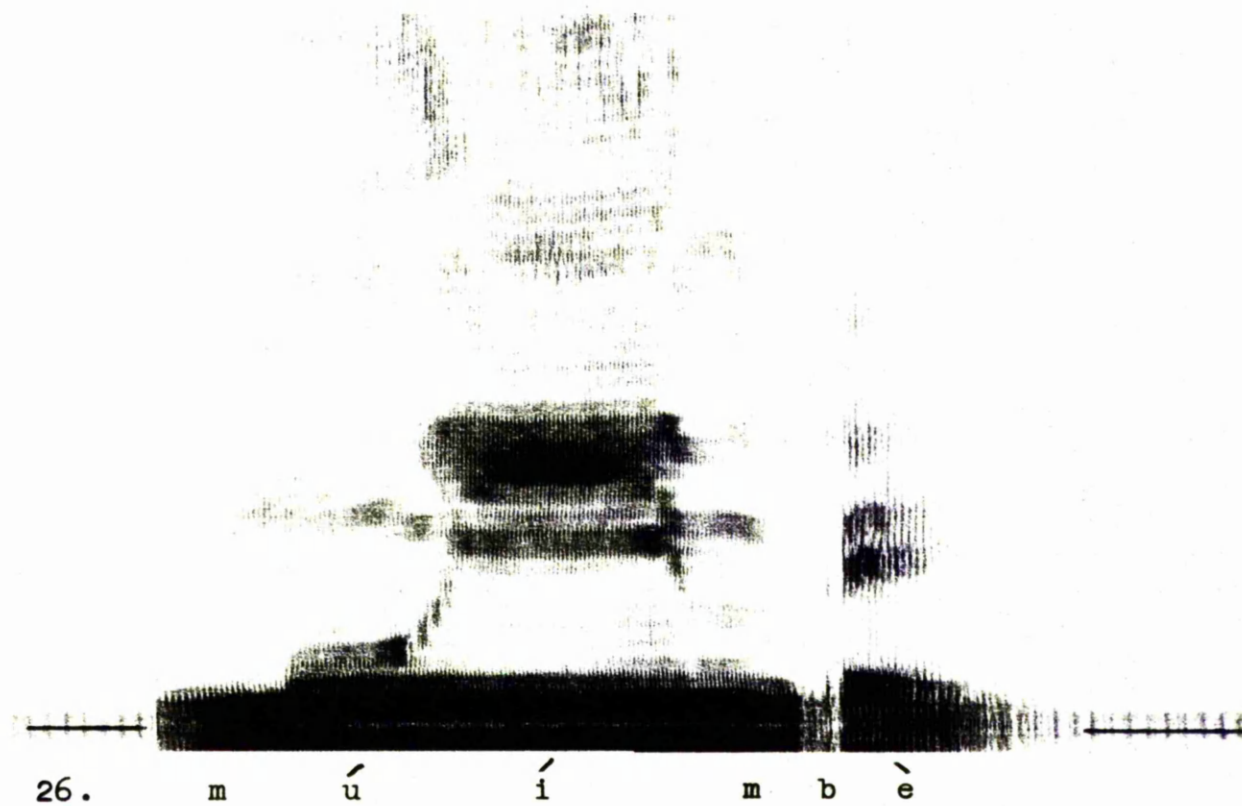
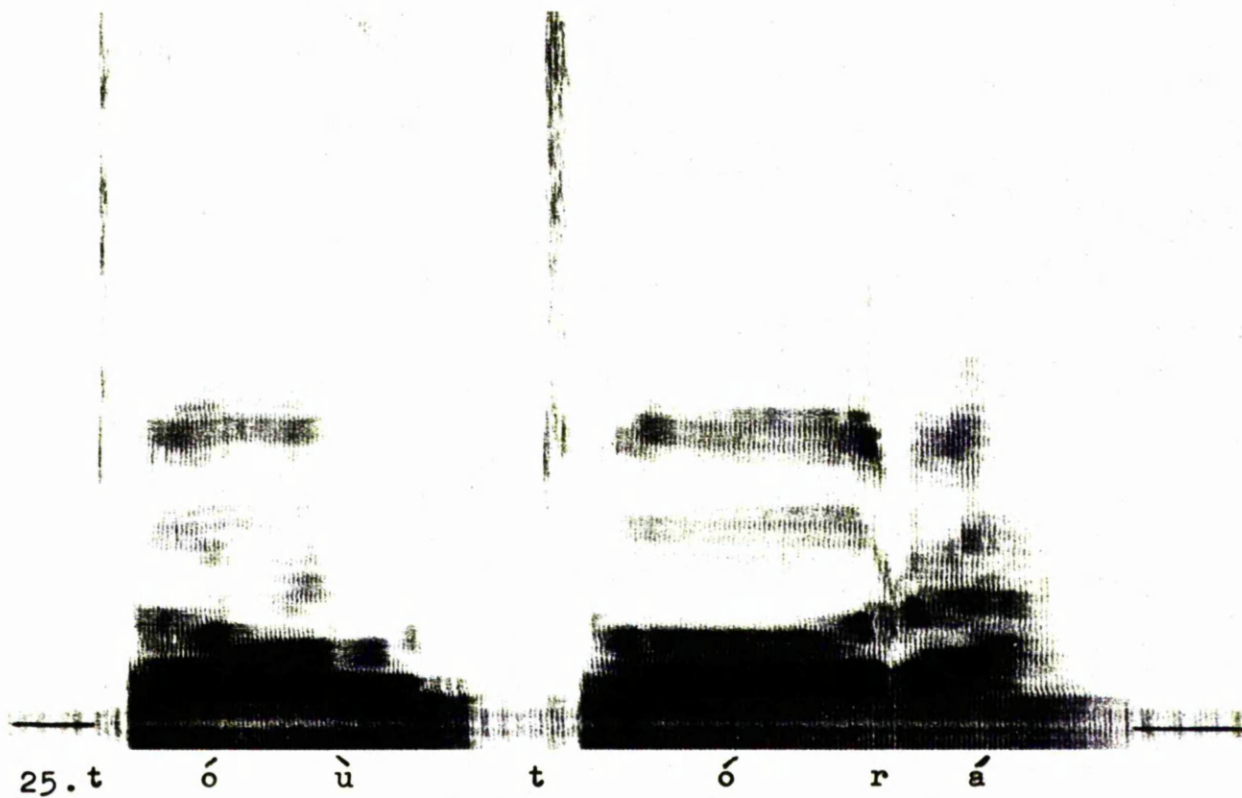
22. t ó í n d à

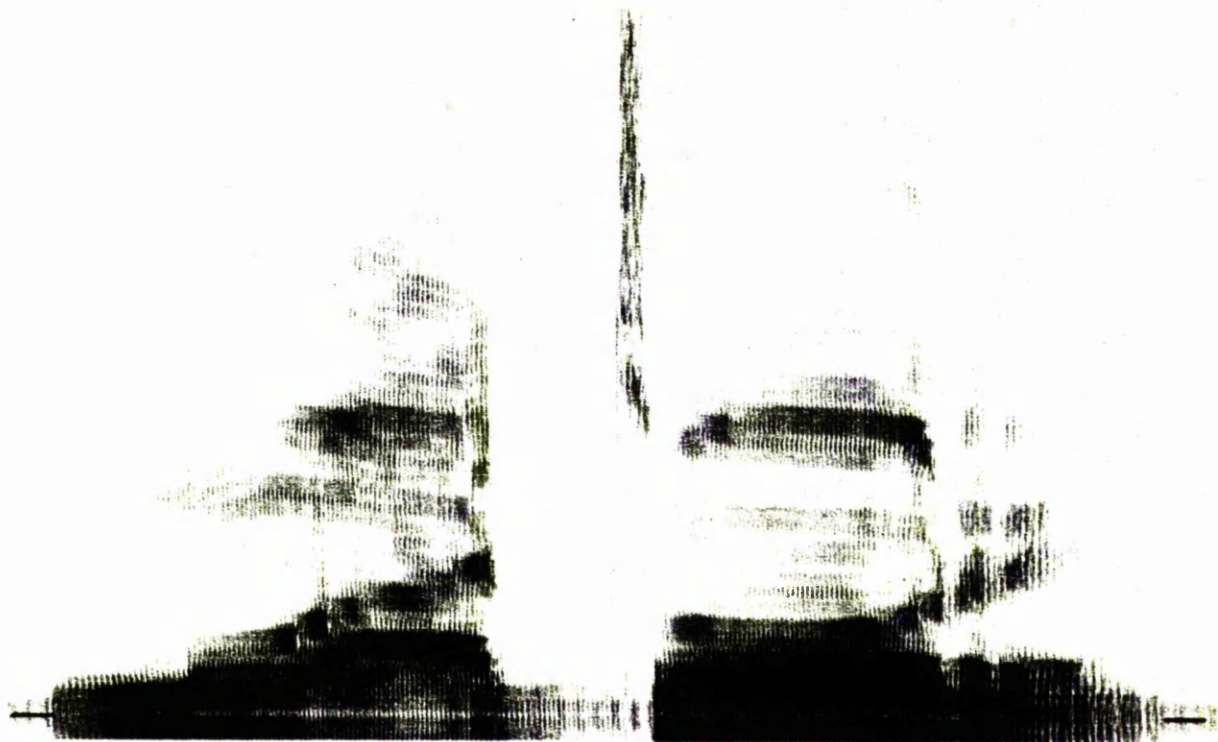


23. t ó é r è r à

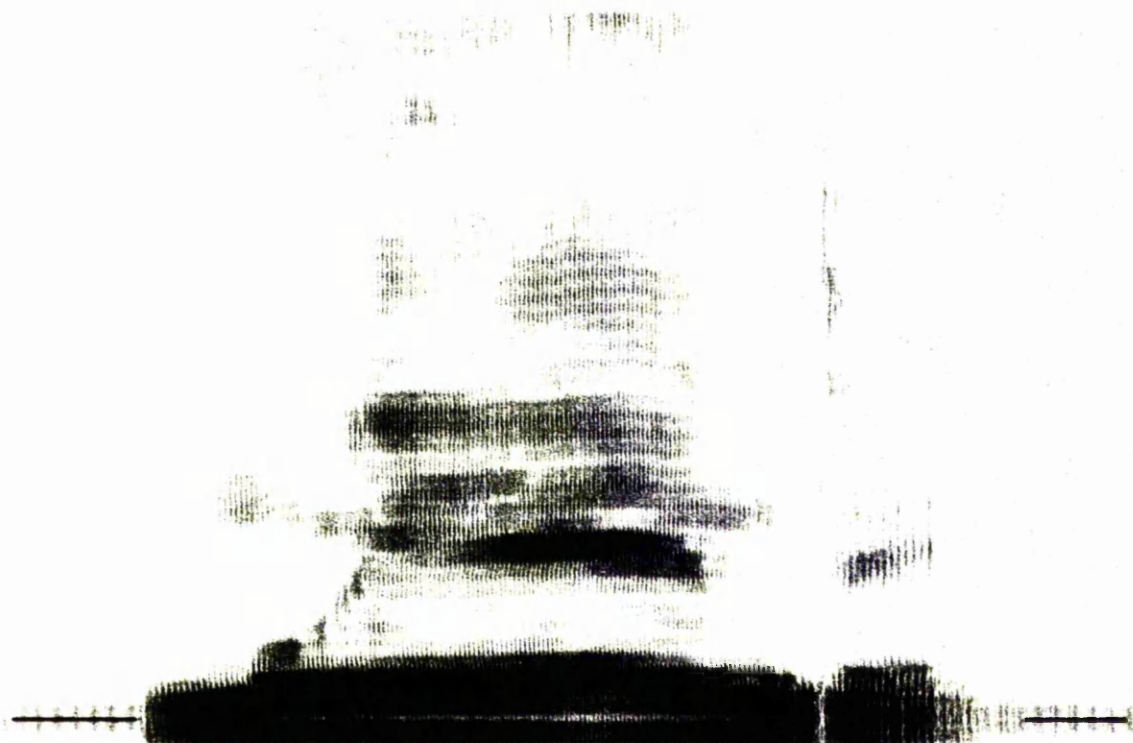


24. t ó à t ó r á

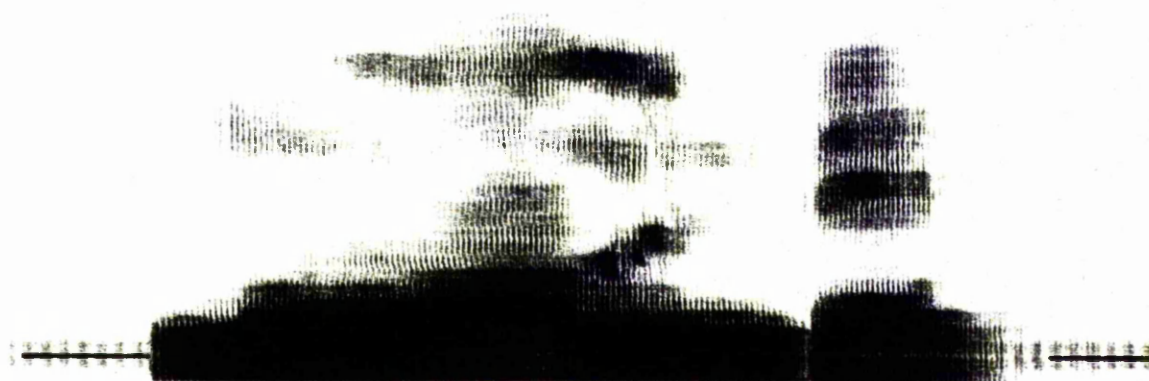




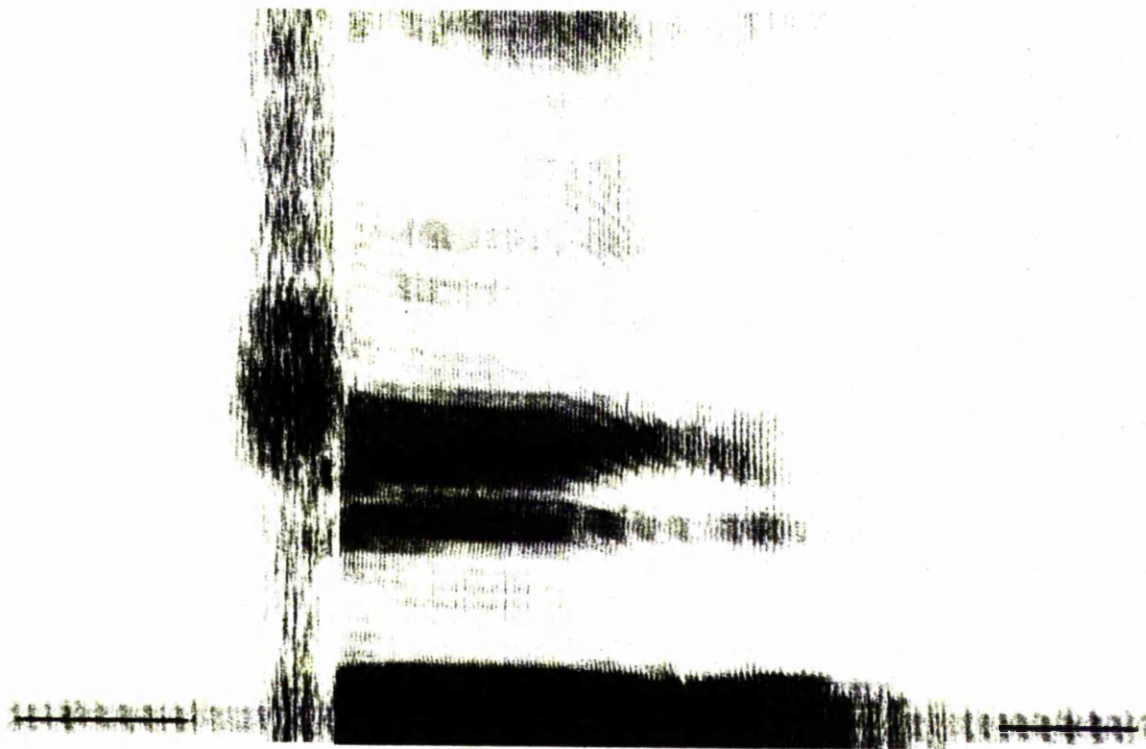
27. m ú à t ó r è



28. m ú é n d è



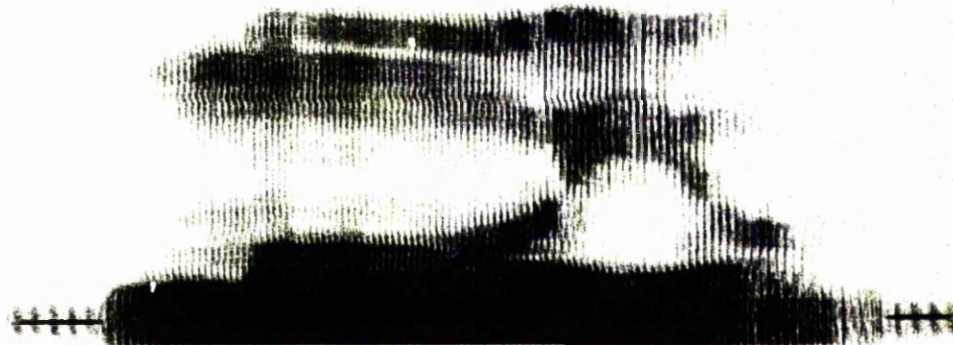
29. m ù ò n d é



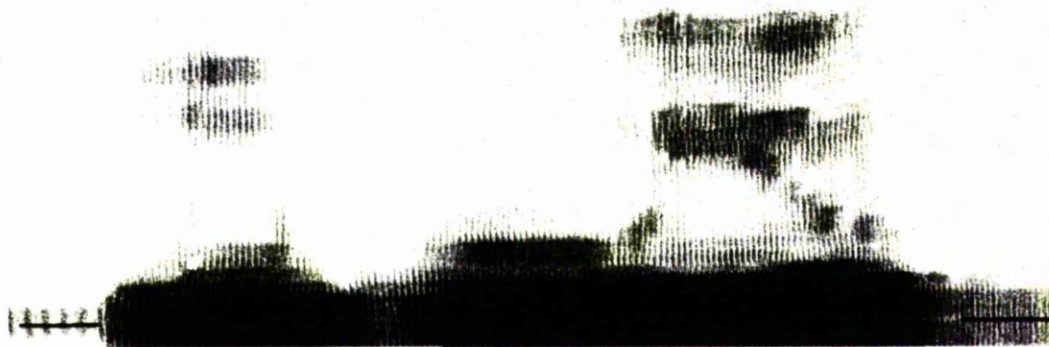
30. t s í j ì



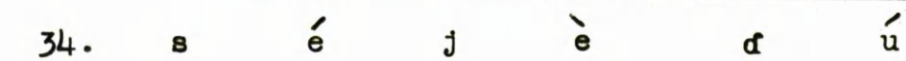
31. á j á m b ù k à



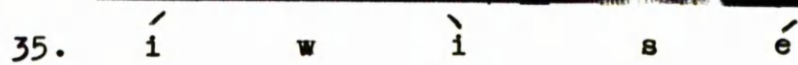
32. m ò j ò



33. m ù v ù j ú



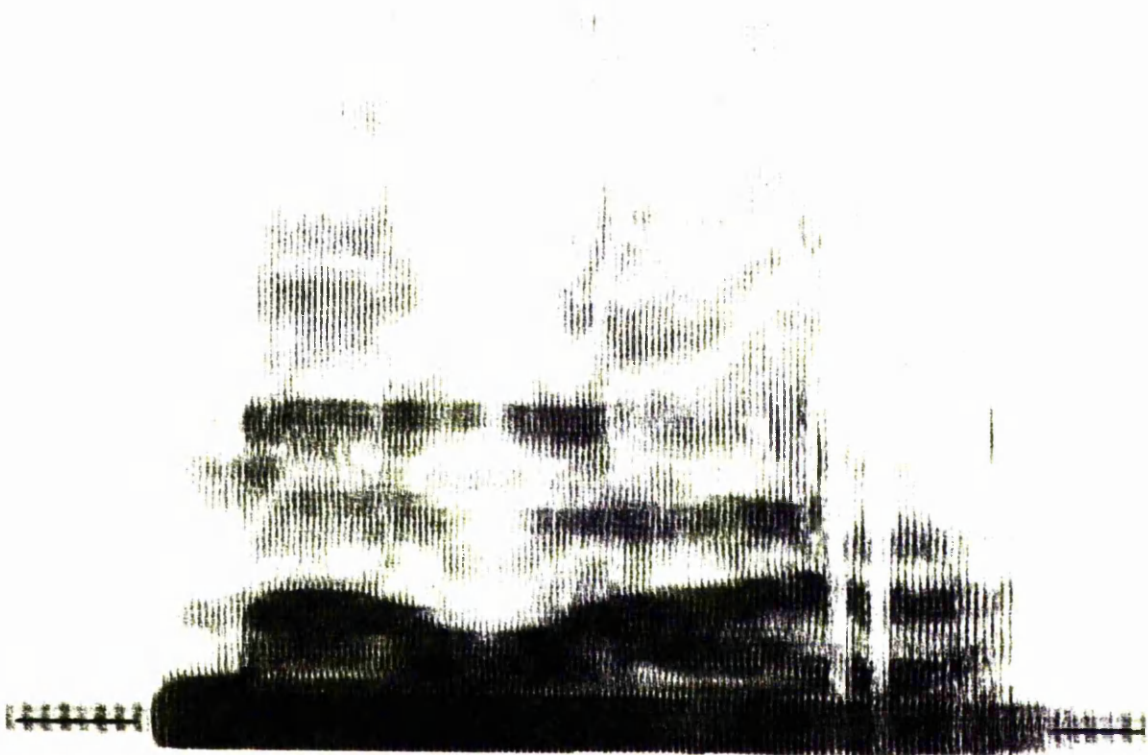
34. s é j è d ú



35. í w ì s é



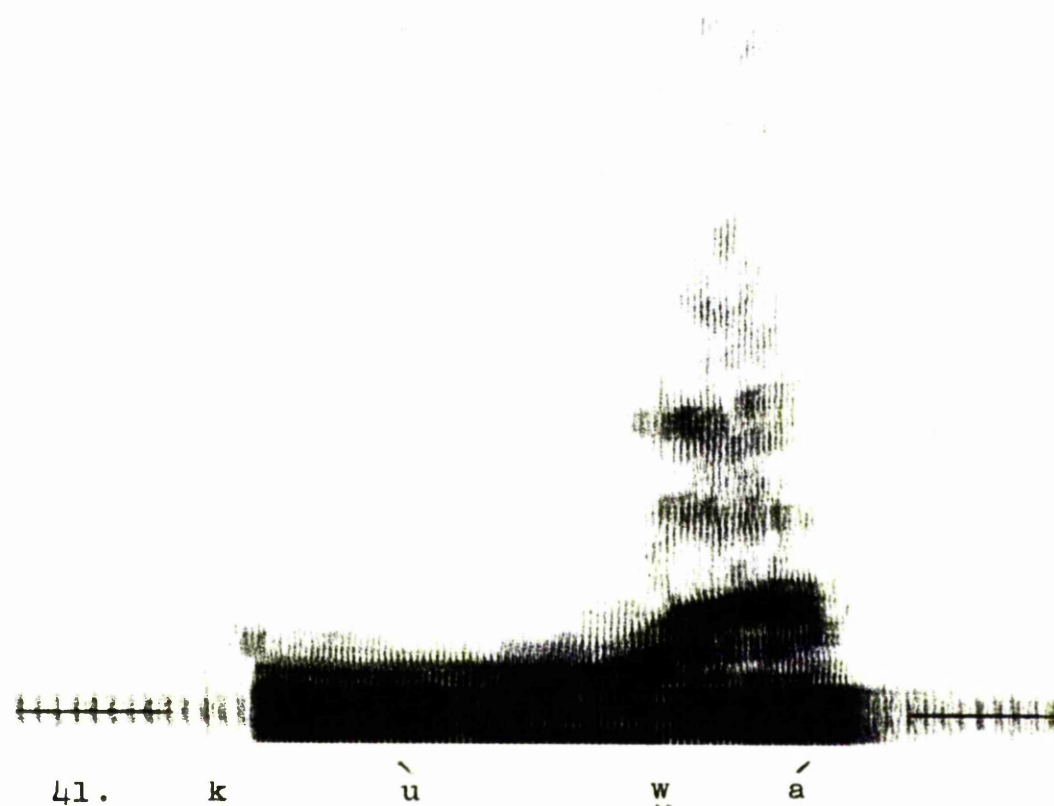
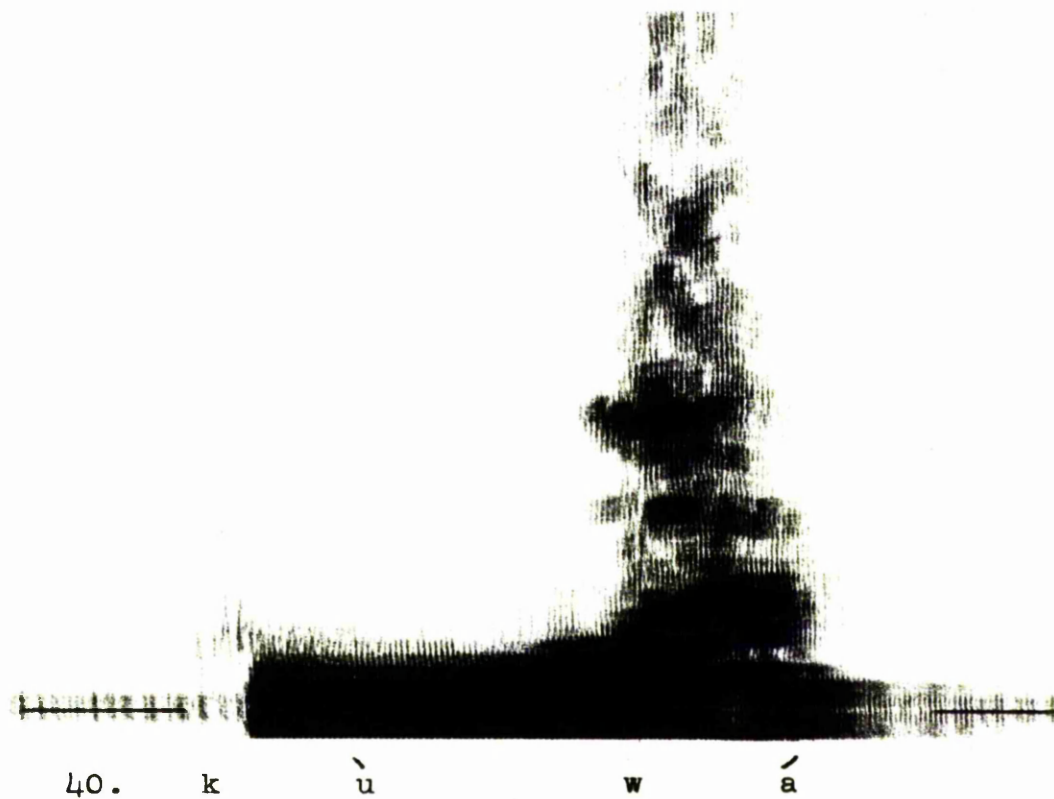
36. s è w é

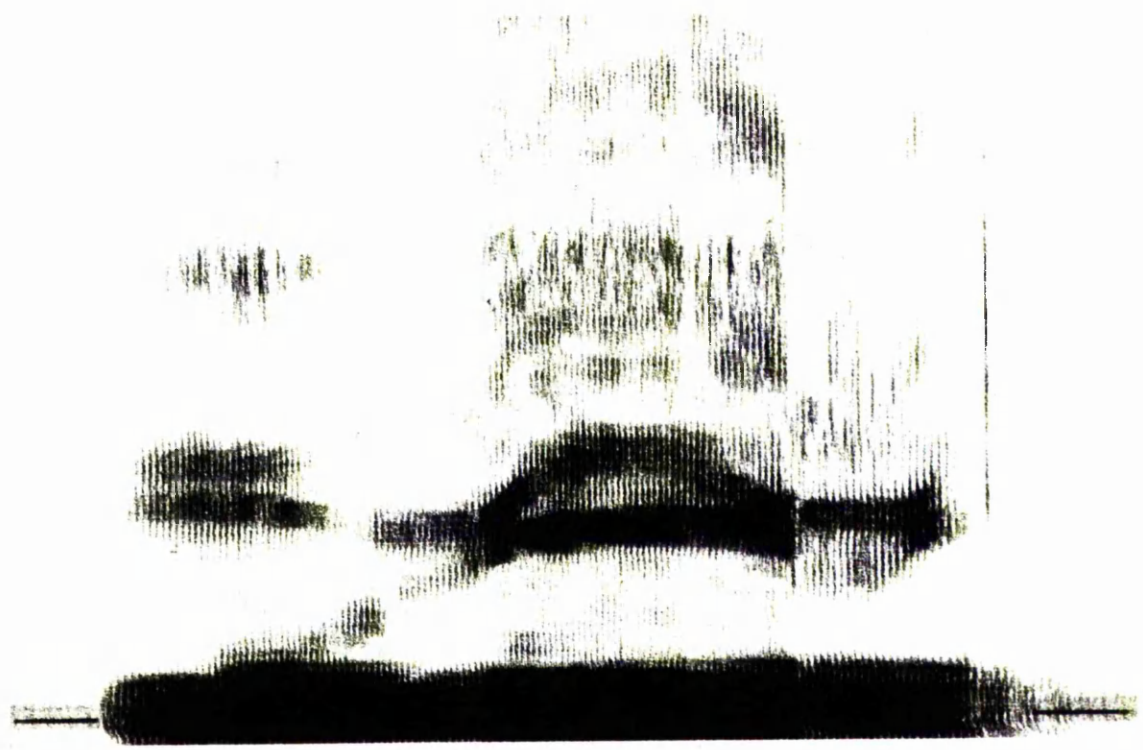


37. m à w à r à

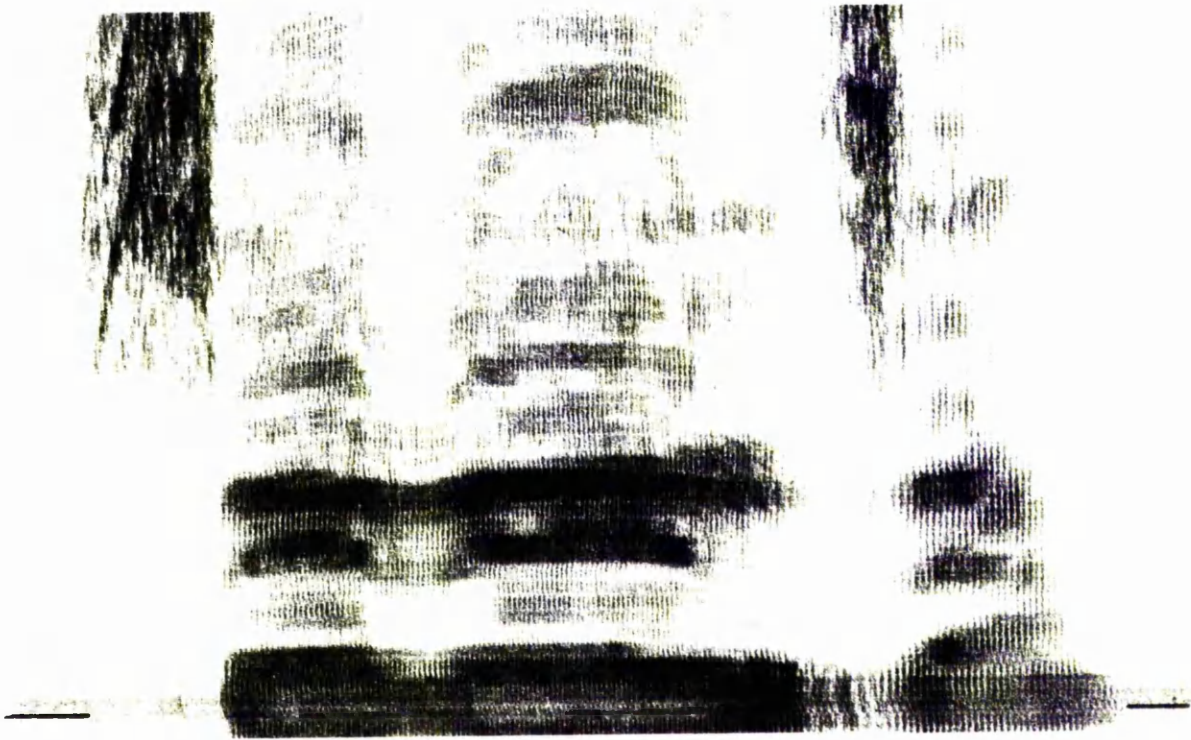
38. m ú w ú j è

39. m ó w ó n è k à

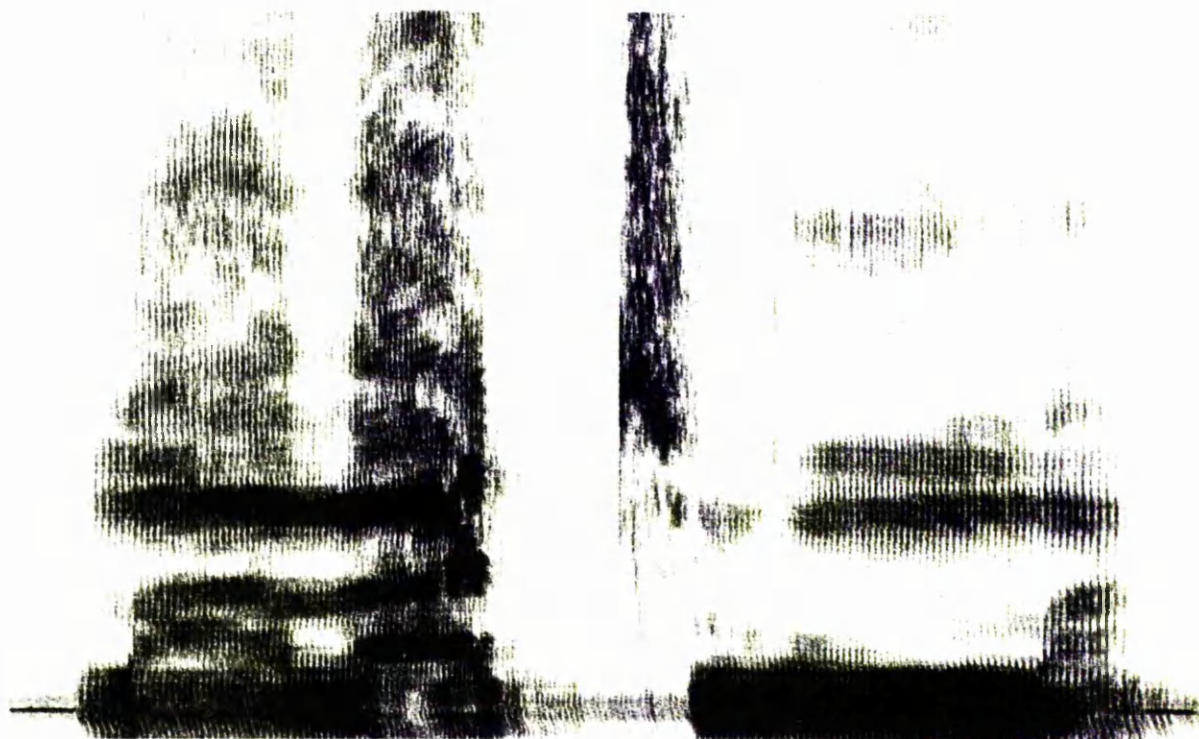




42. m ù v ì r ì



43. s é v é ñ z á



44. m à v á t ù m à



45. m ù v ù j ú



46. s o v o p e n g a



47.

b.

i

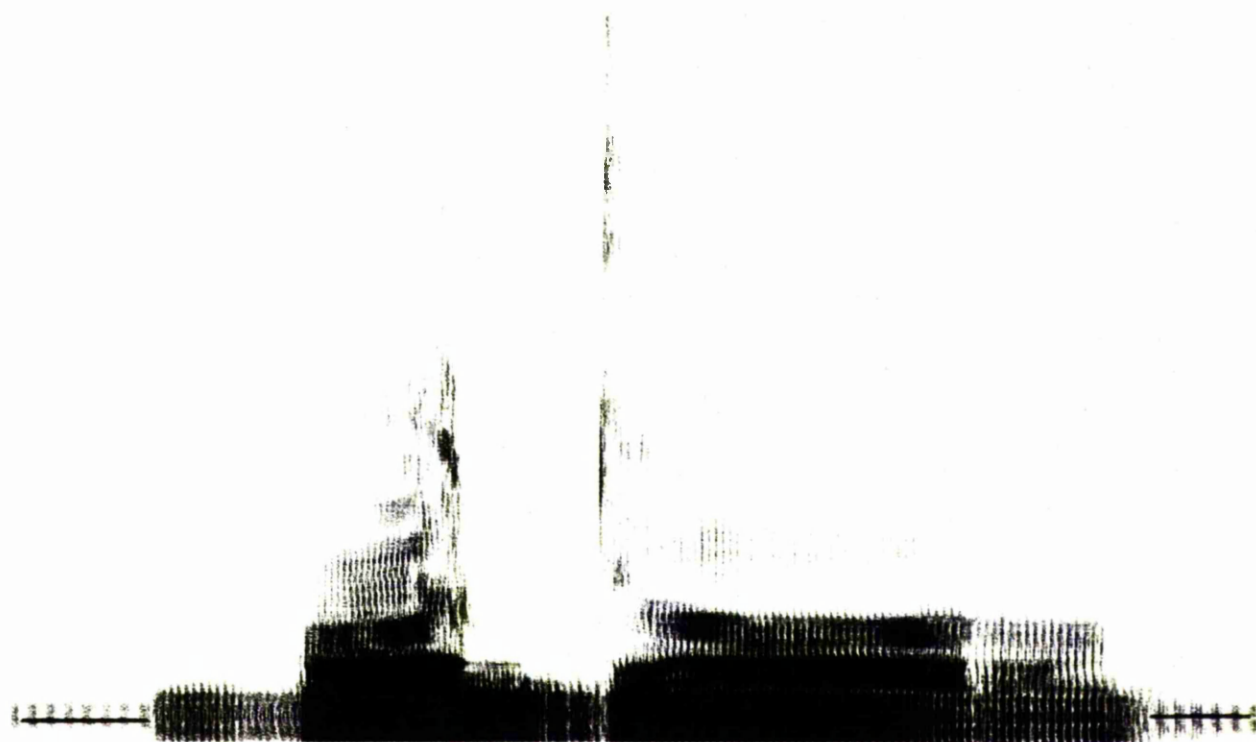
z.

a



47.

[b] Section



48.

b

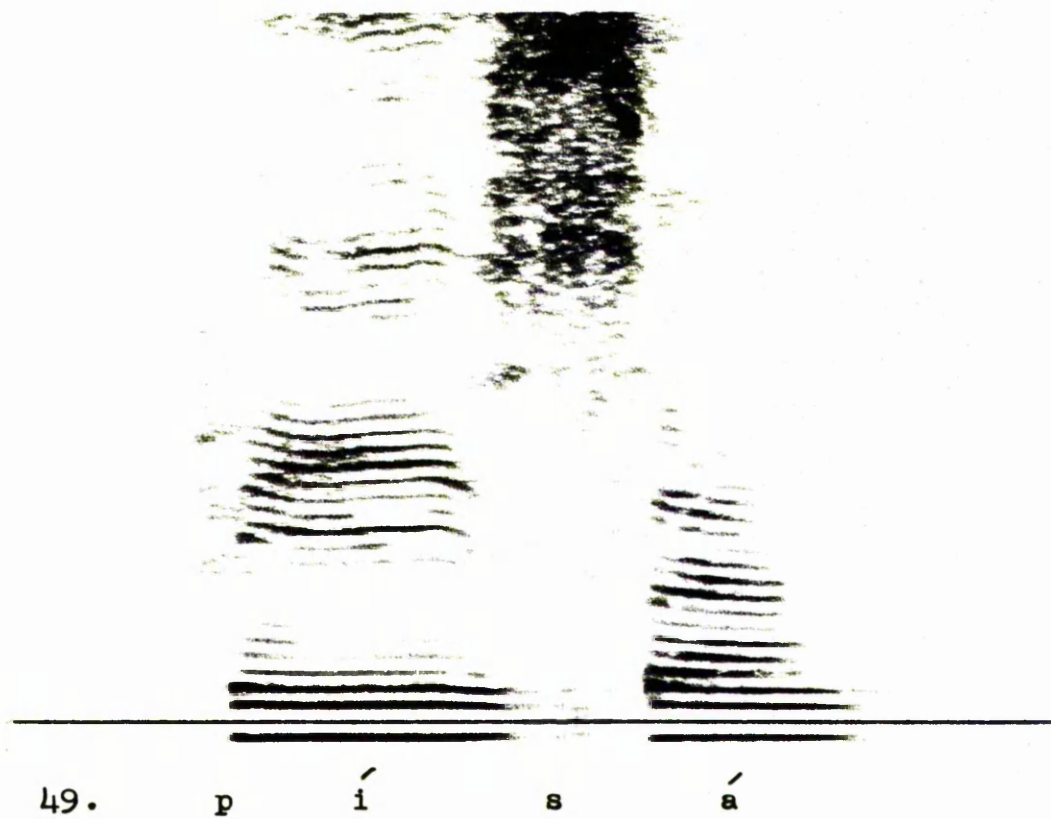
ó

ð

ò

r

ò



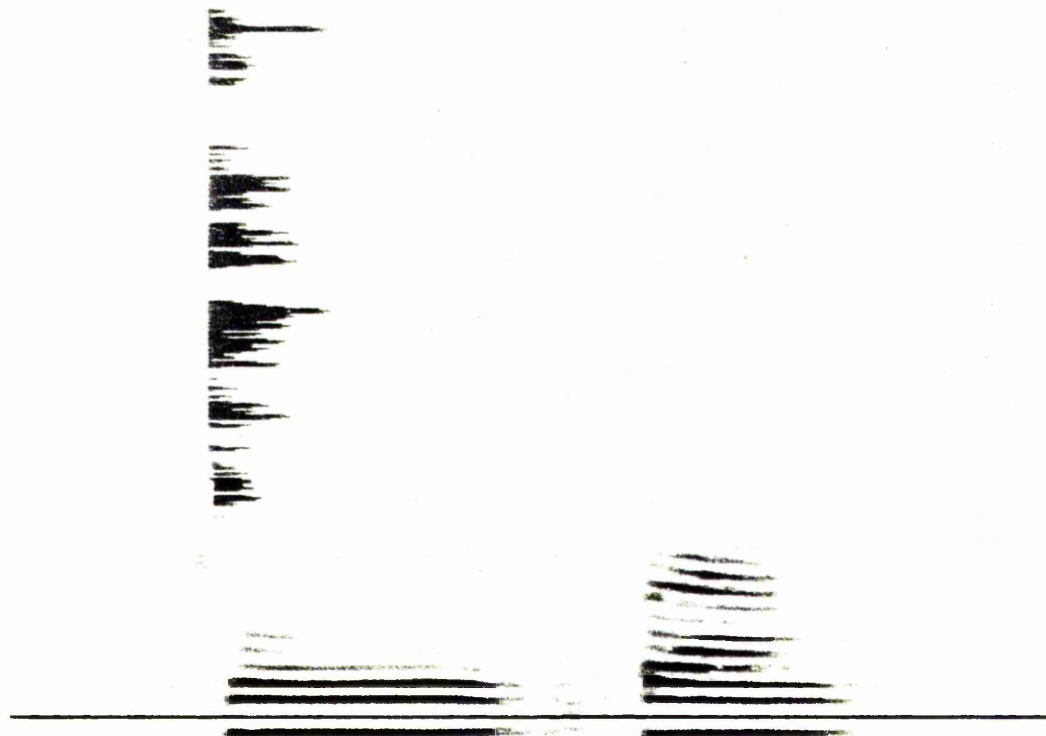
49.

p

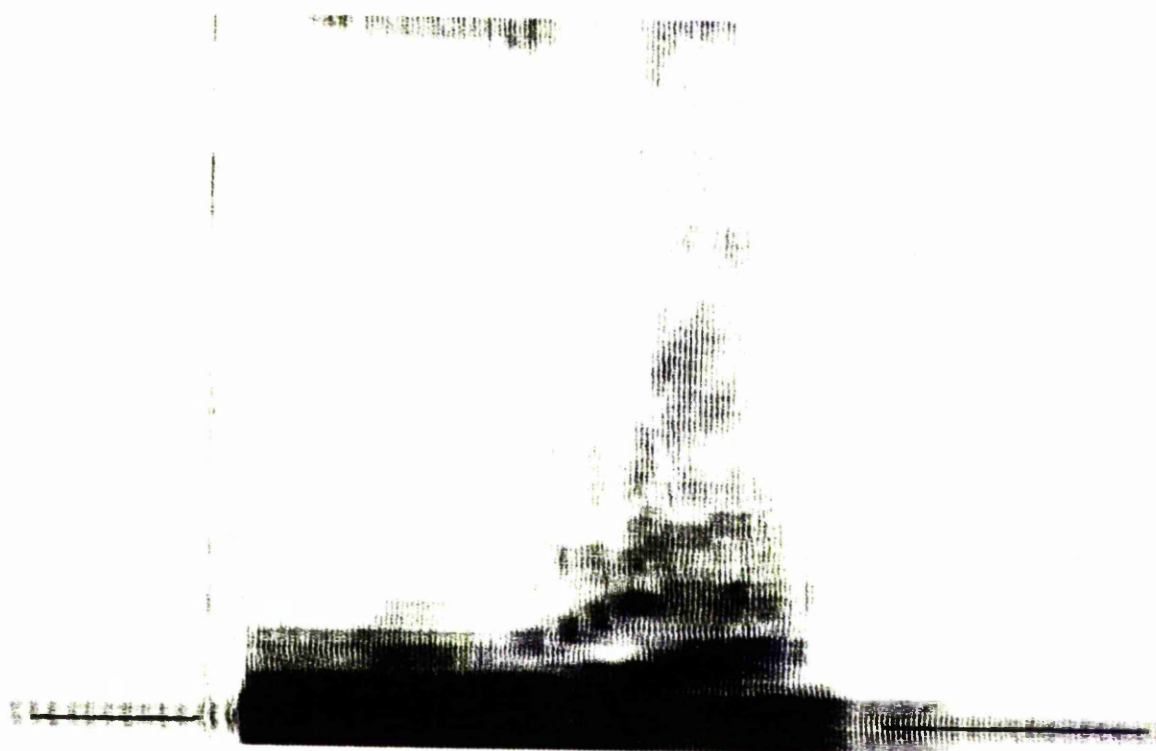
í

s

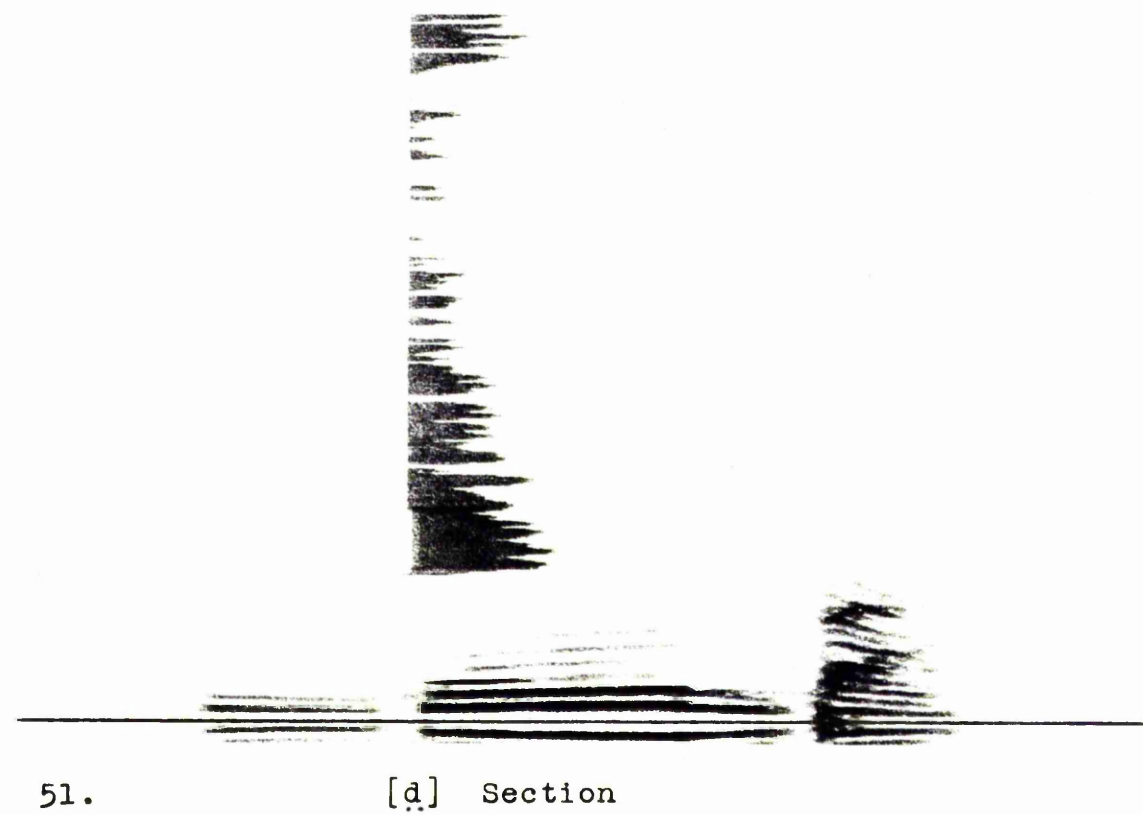
á



49. [p] Section



50. p ú r á





52.

à

ò

r

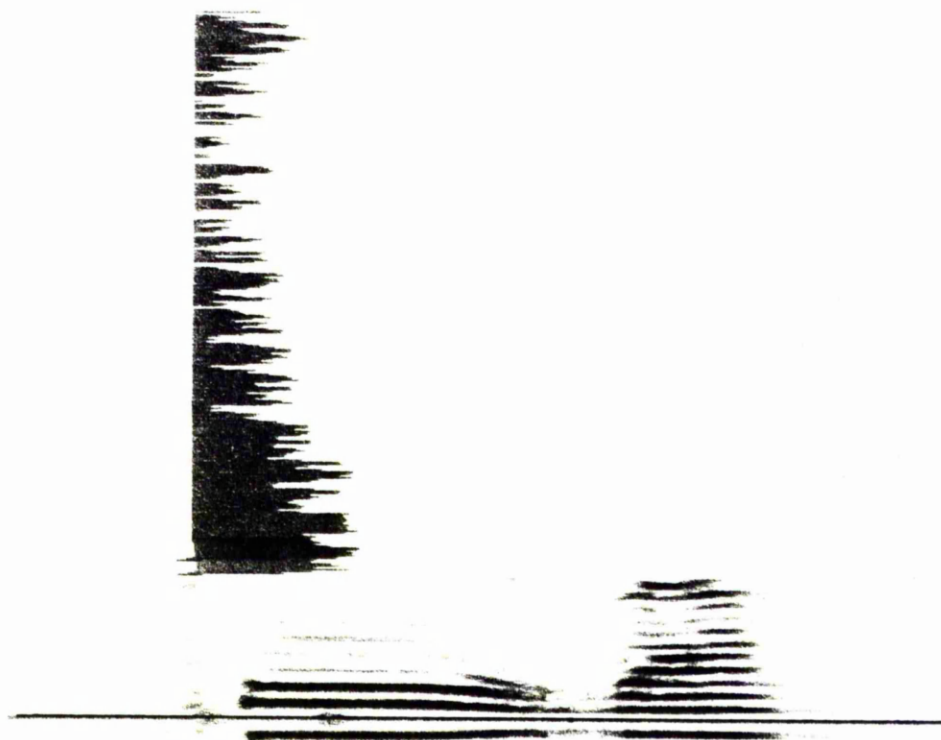
ó

b

à

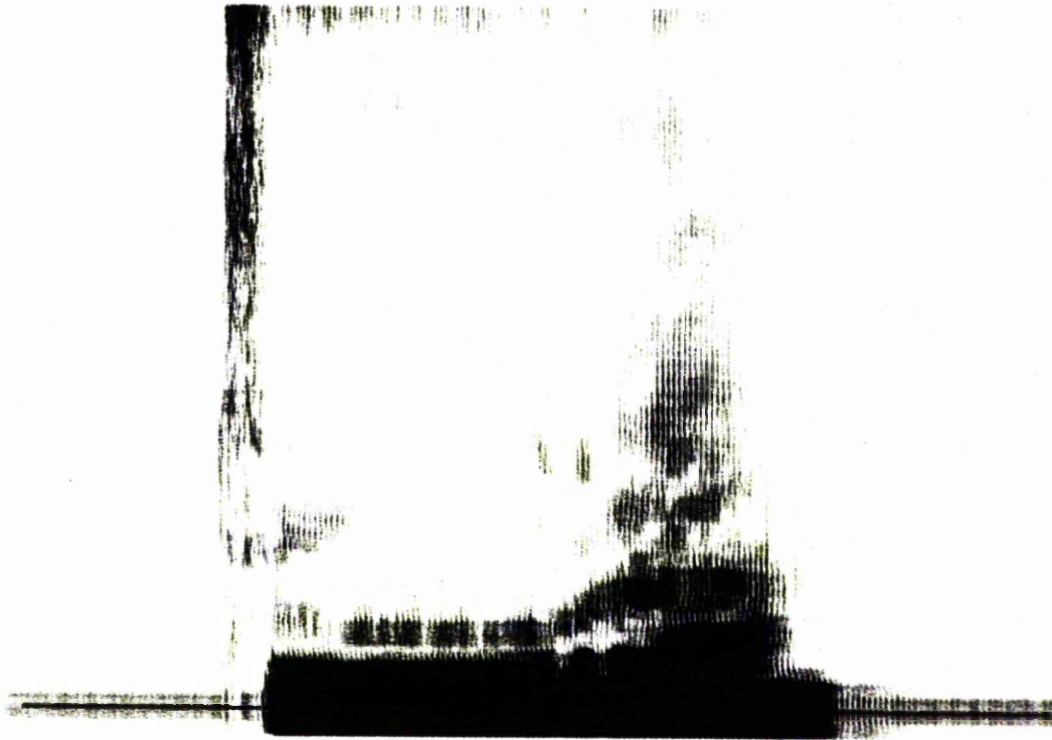


53.



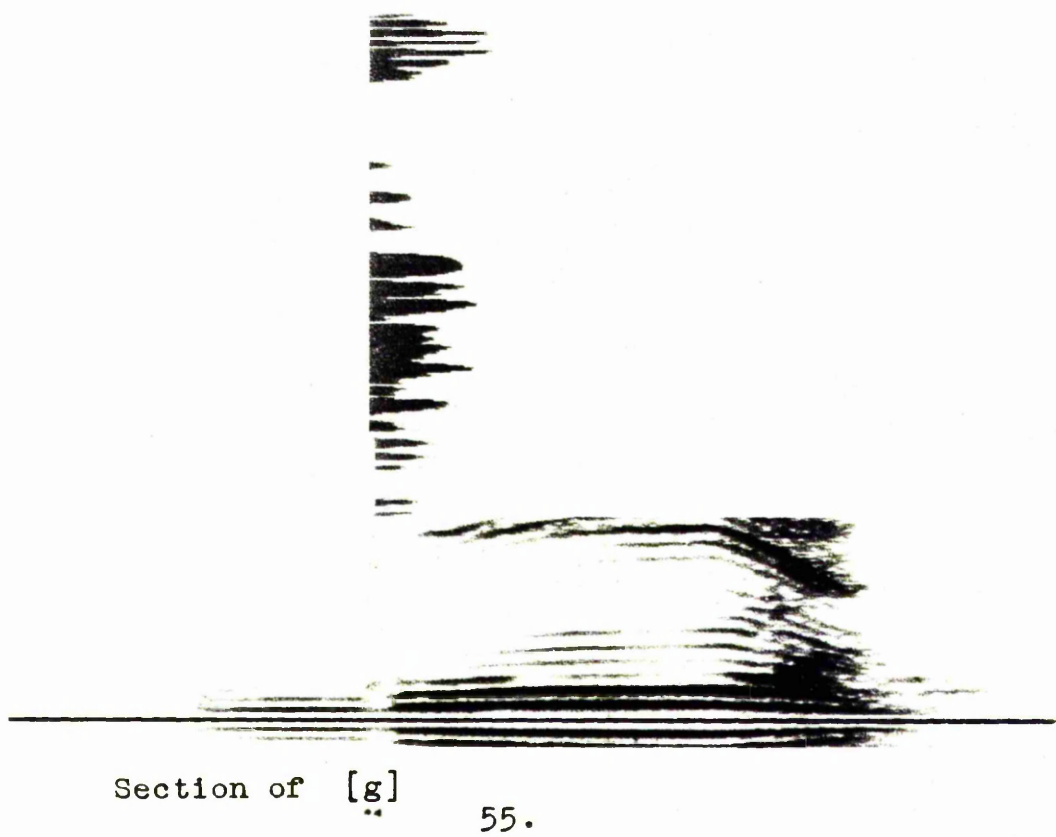
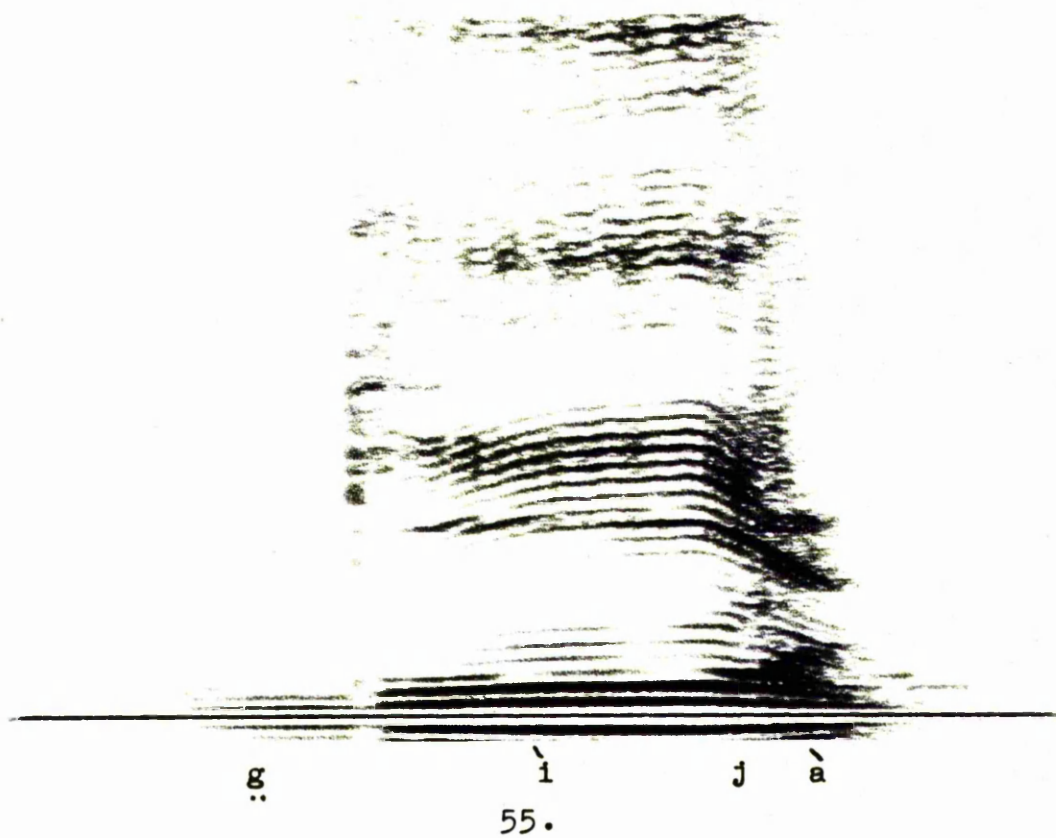
[t] Section

53.



t ú r á

54.



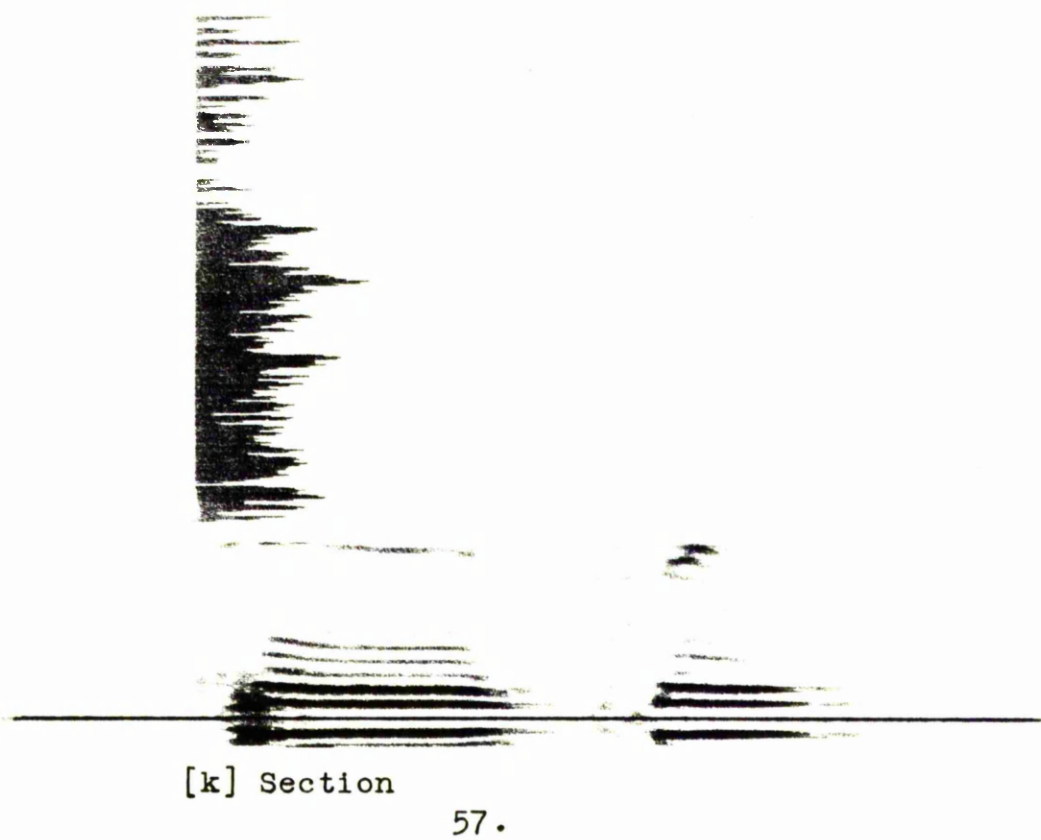
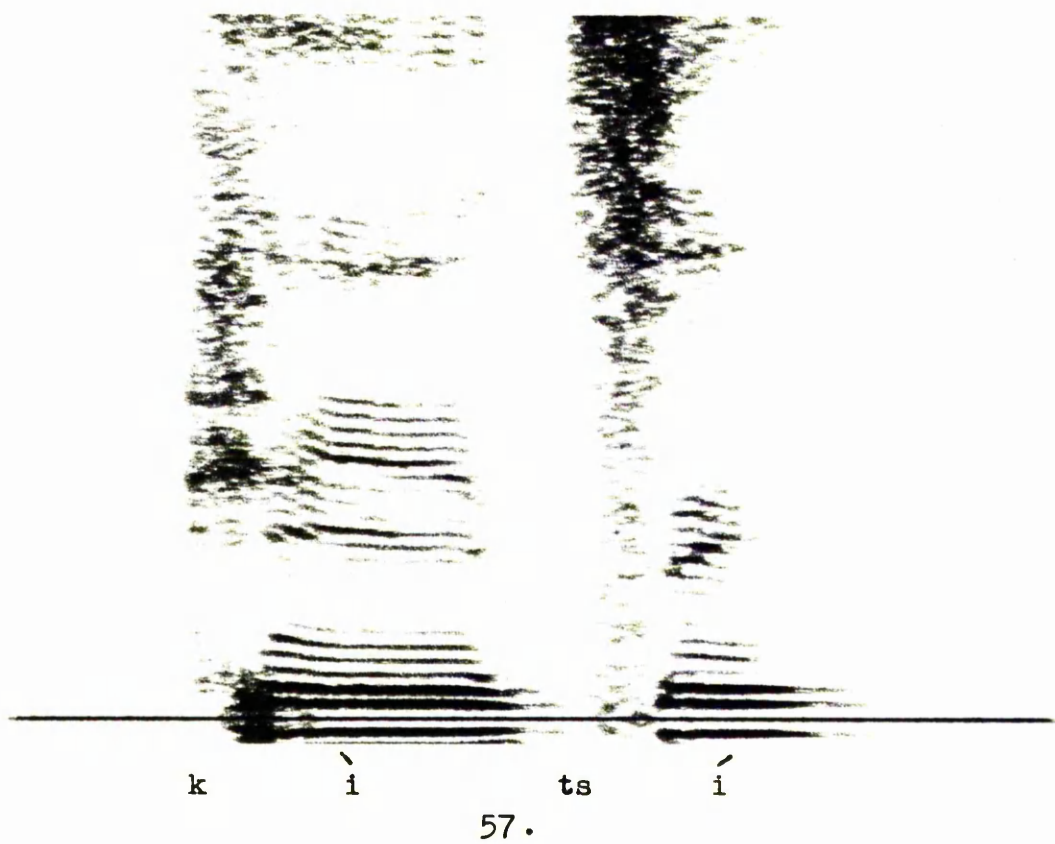


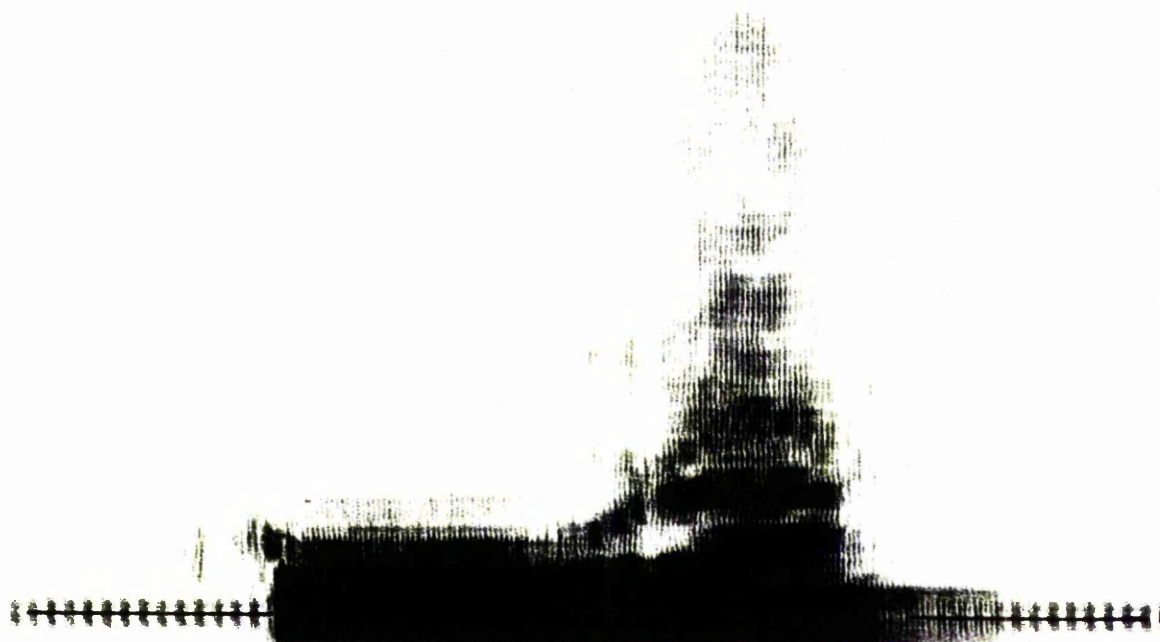
g

ú

r

ú





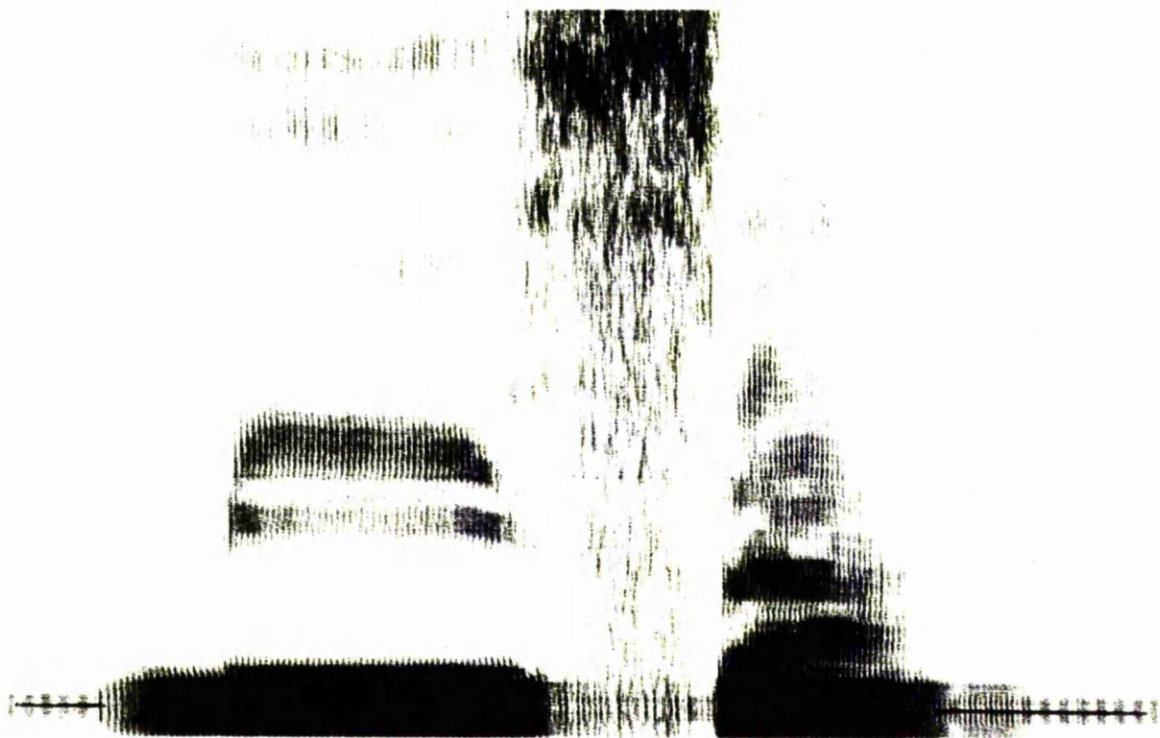
k

ú

r

á

58.



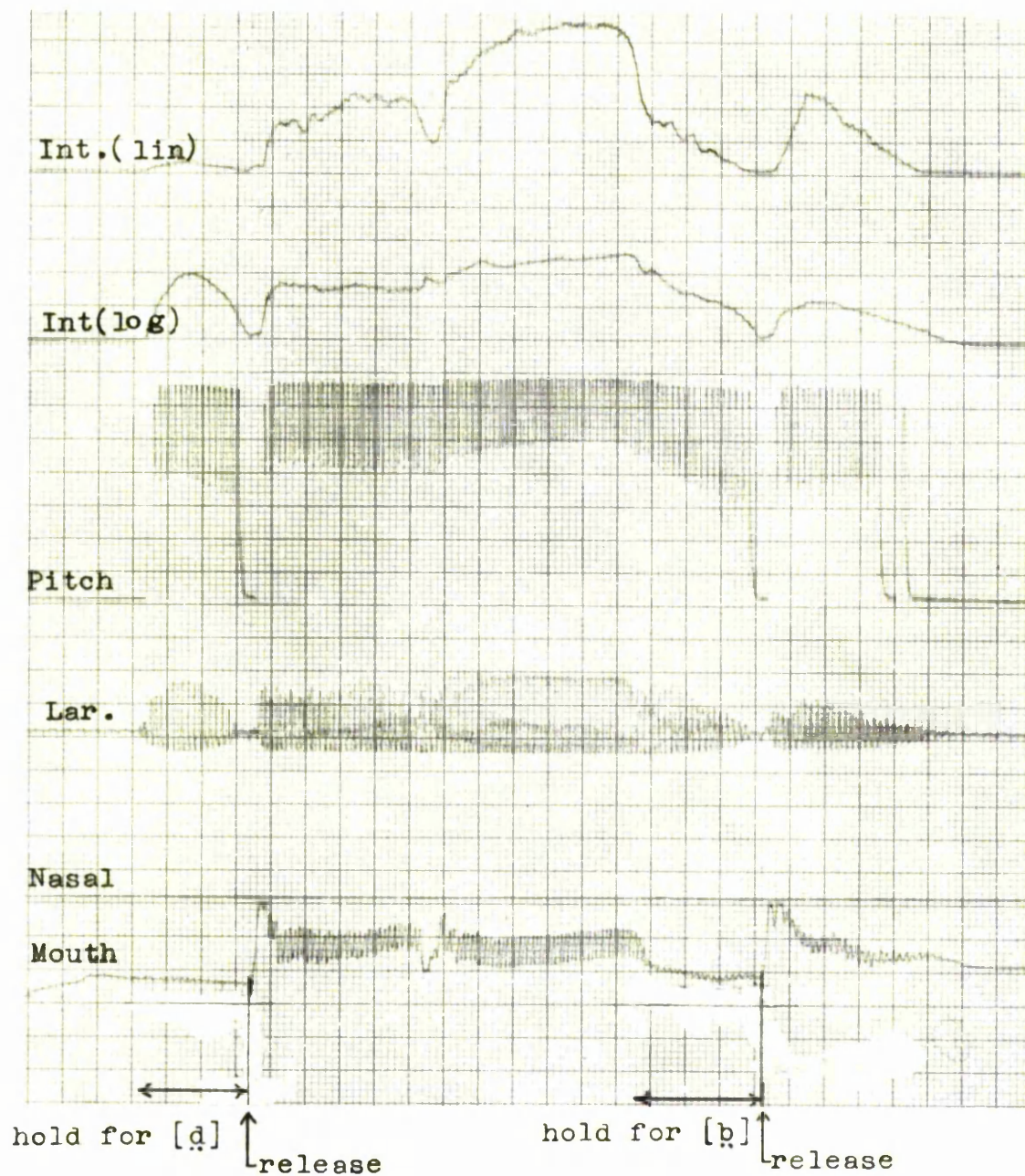
6

i

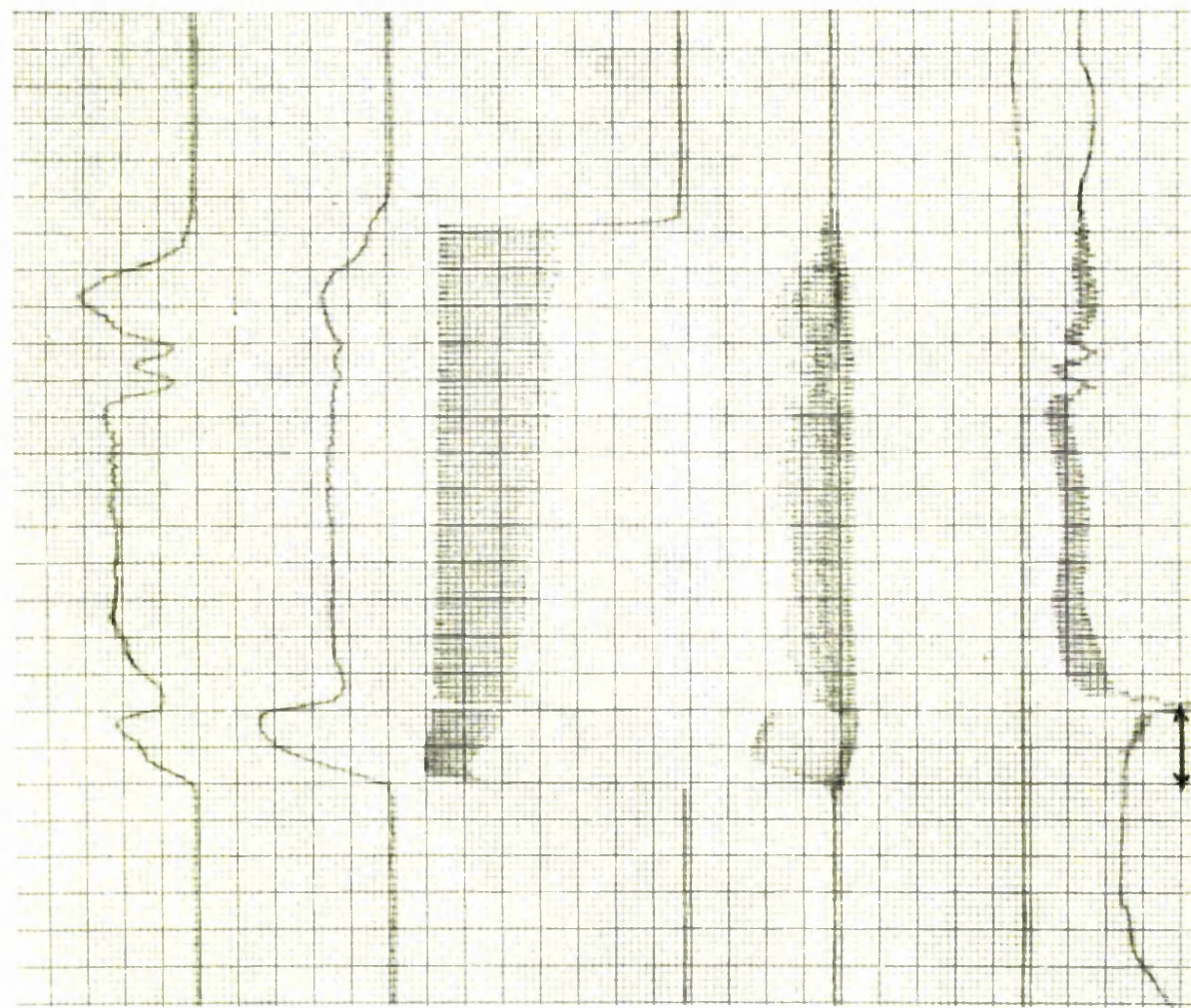
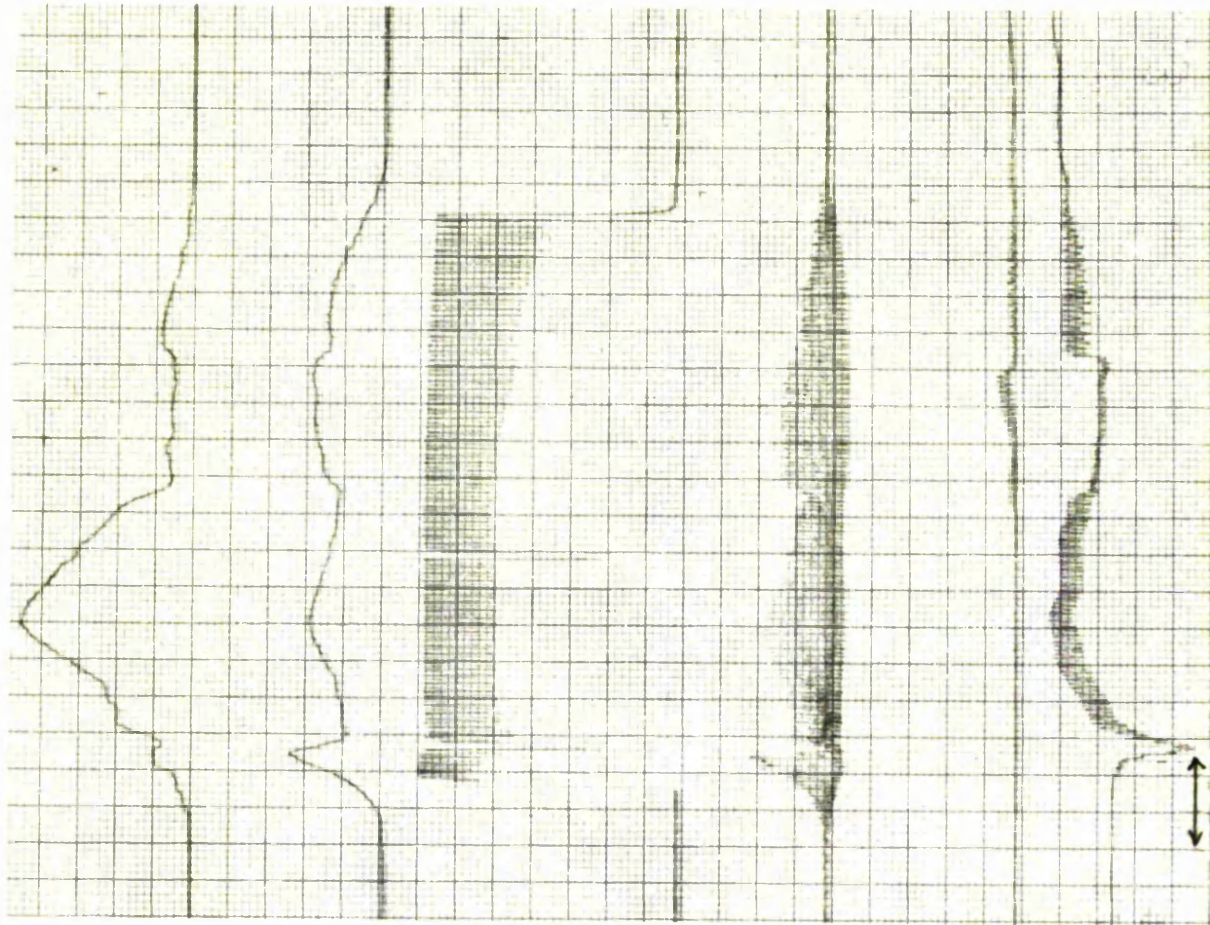
s

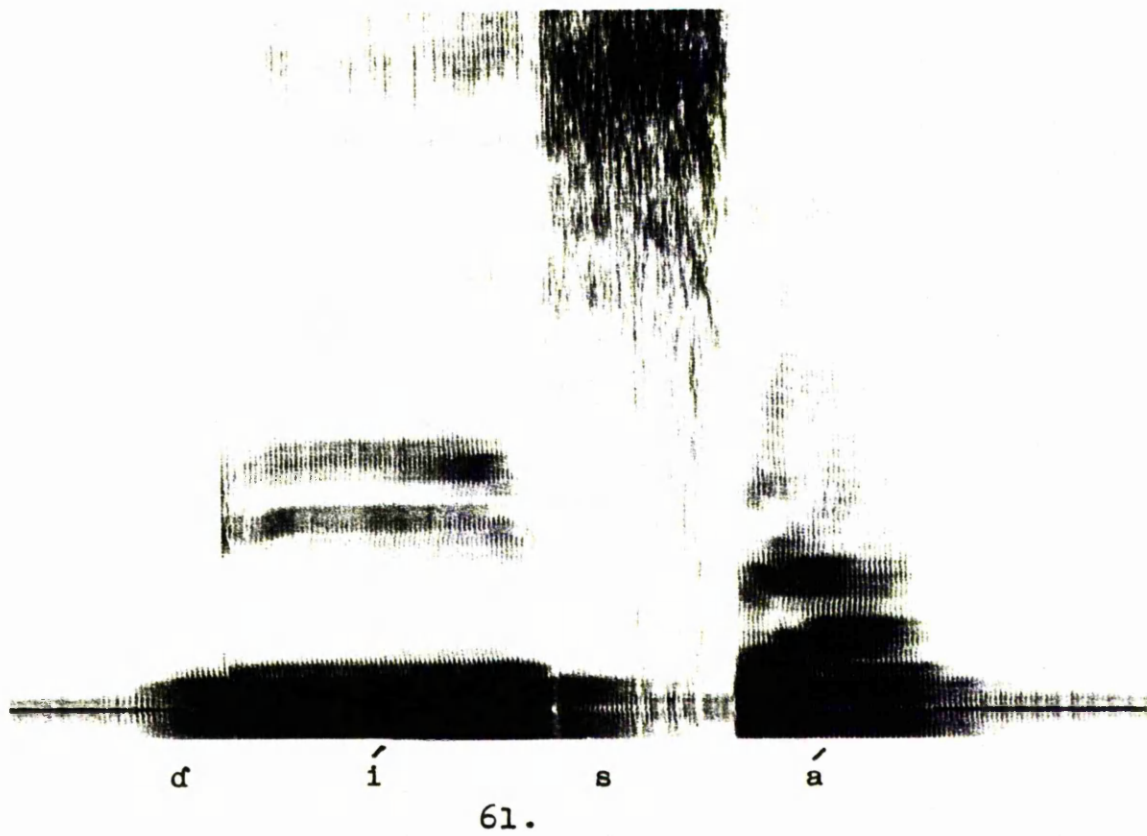
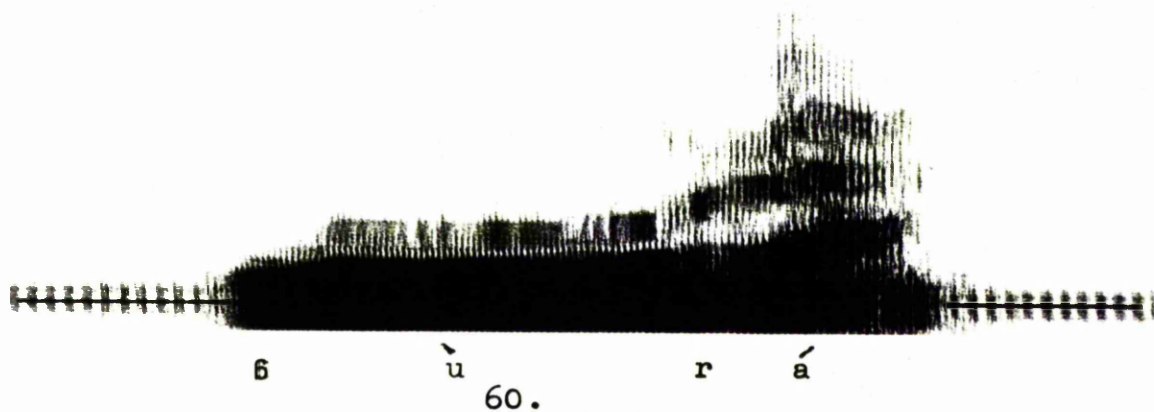
a

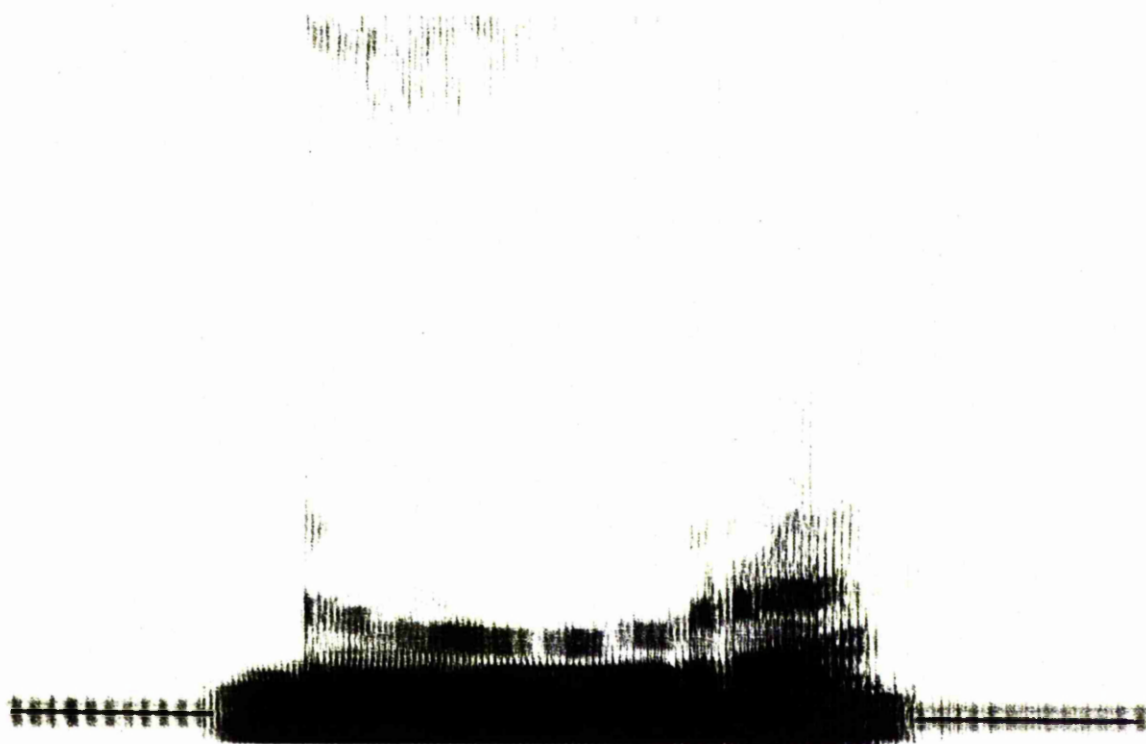
59.



Mingogram 1. [dòróbà] "town"



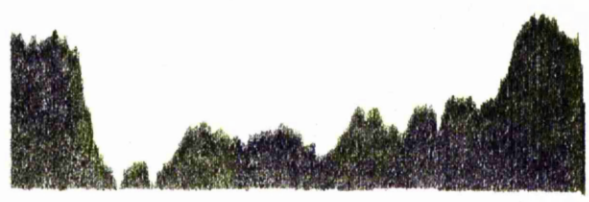




đ ù r á

62.

í s á p á m à r ú dz í
63.



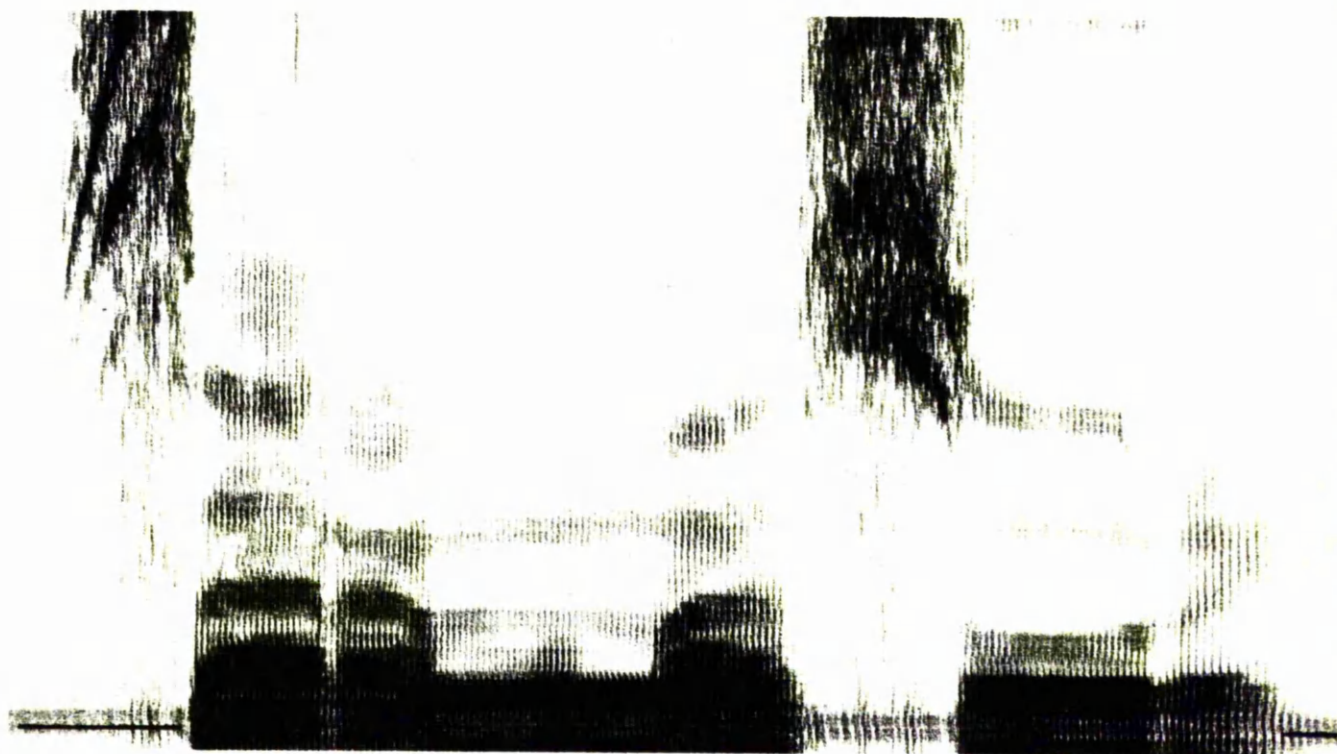
[s] . . . Sections of [r]
63.

p á n è m á v ù dz ì

64.

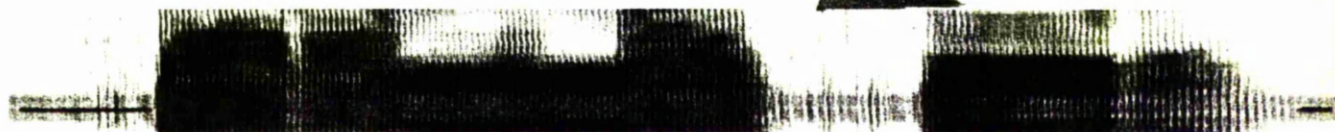
Section of [v]

64.



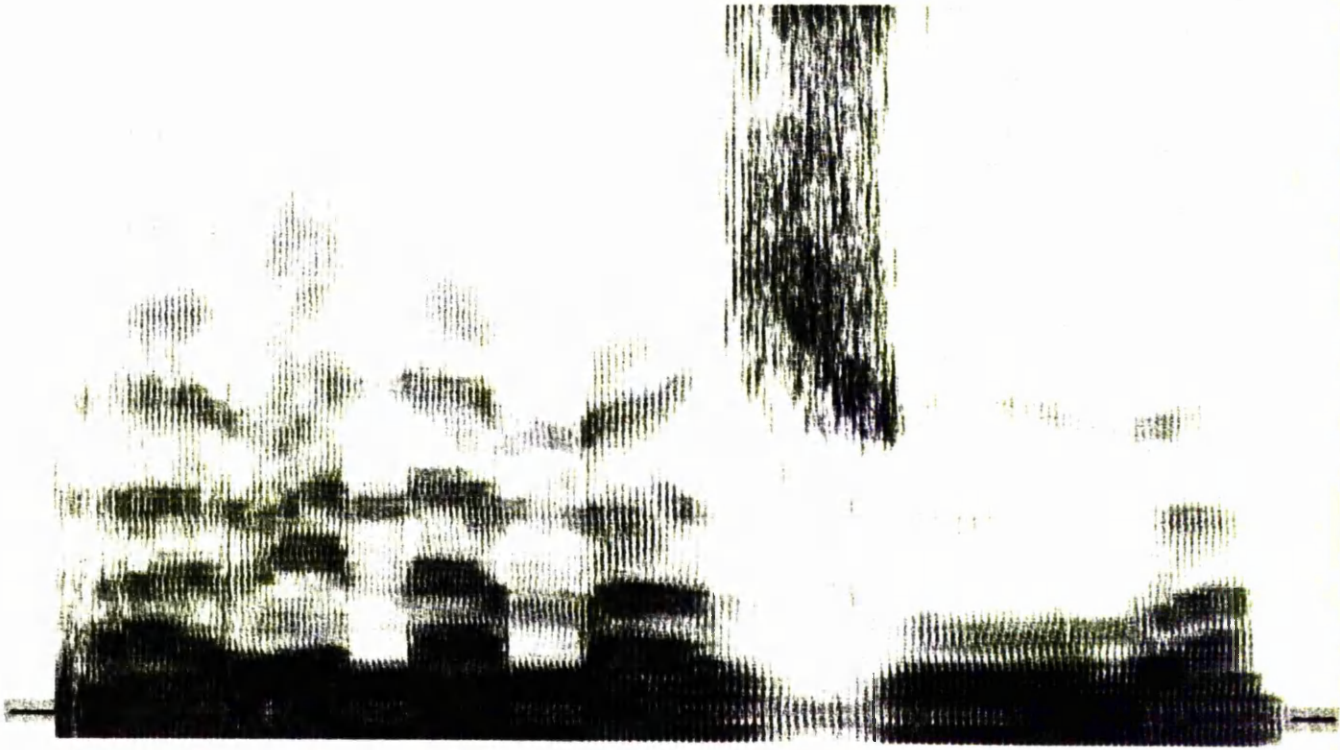
s á r á m ú m à s ú r è

65.



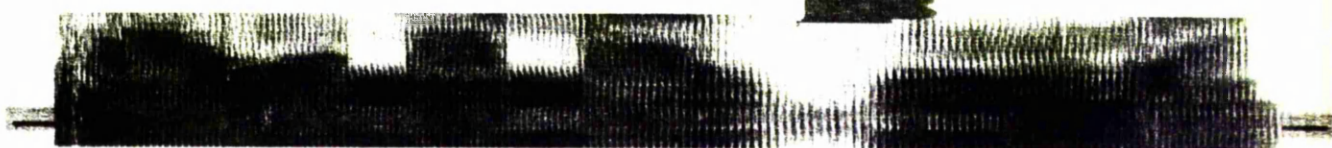
Section of [s]

65.

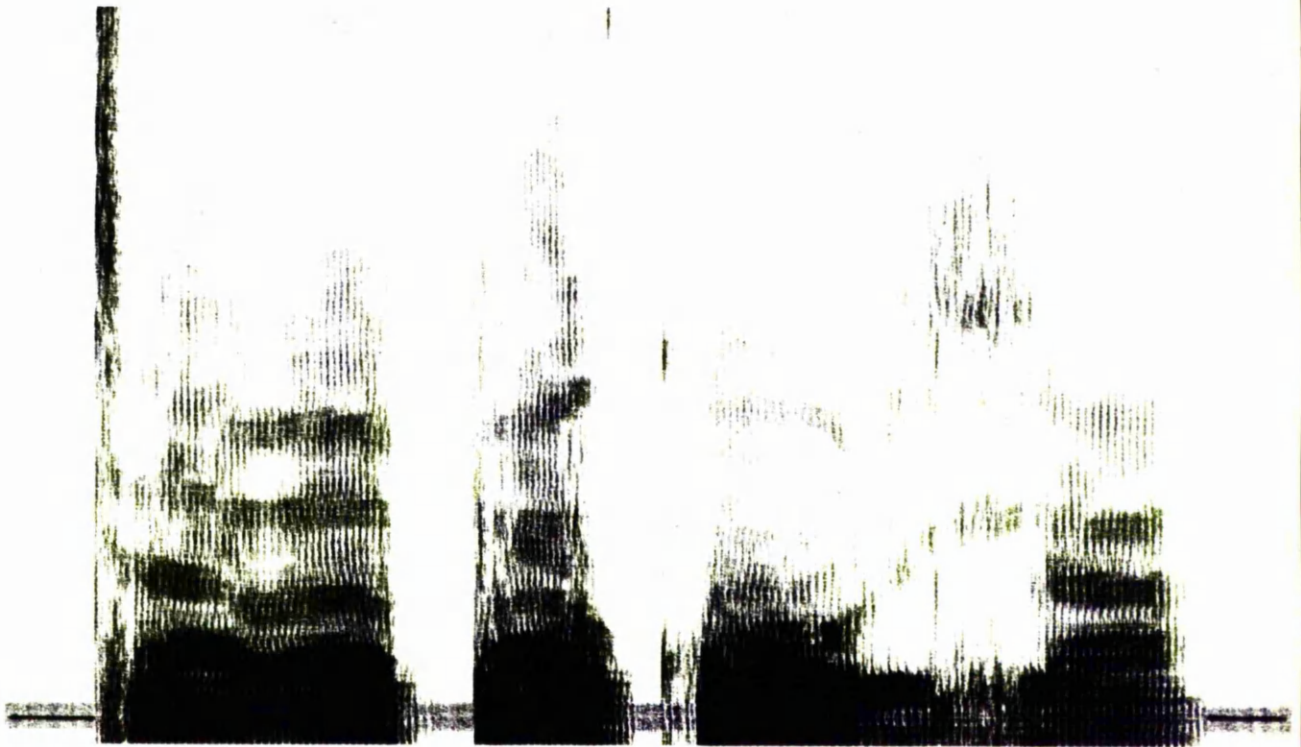


p á v é n á m à z ú v á

66.

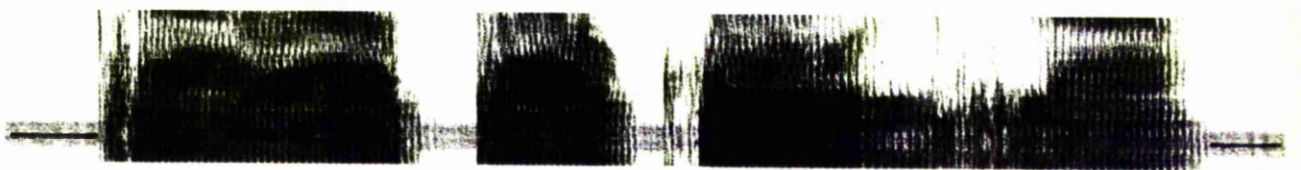


Section of [z]
66.



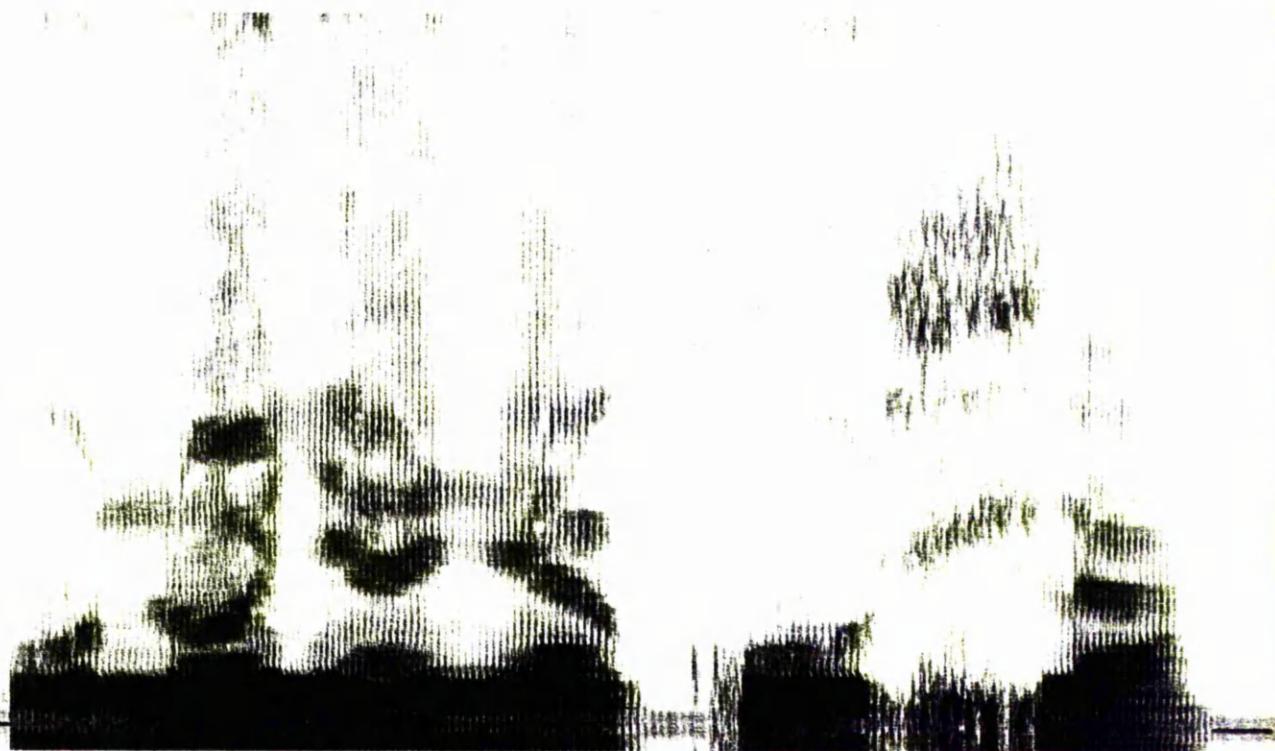
t à v á p á k ù h á

67.



Section of [h]

67.



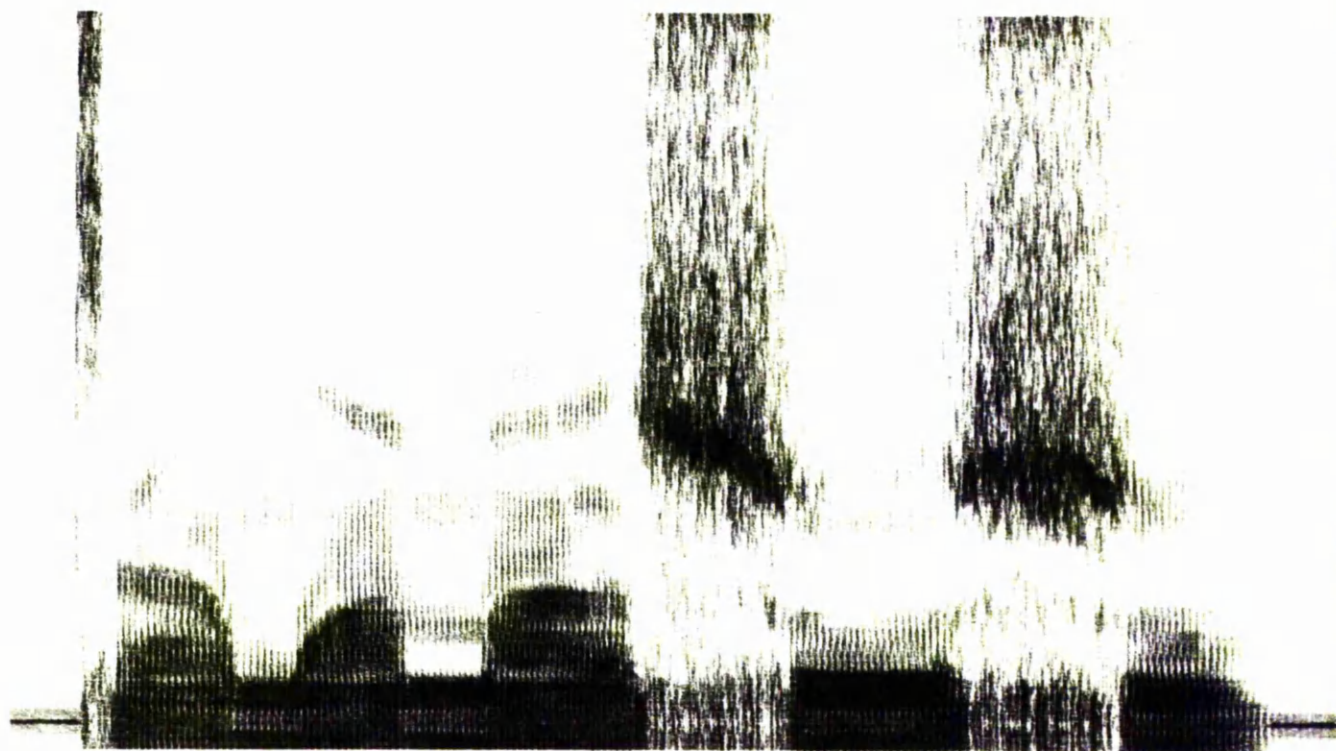
ù n ó ò à ò à k ù é á

68.



Section of [4]

68.



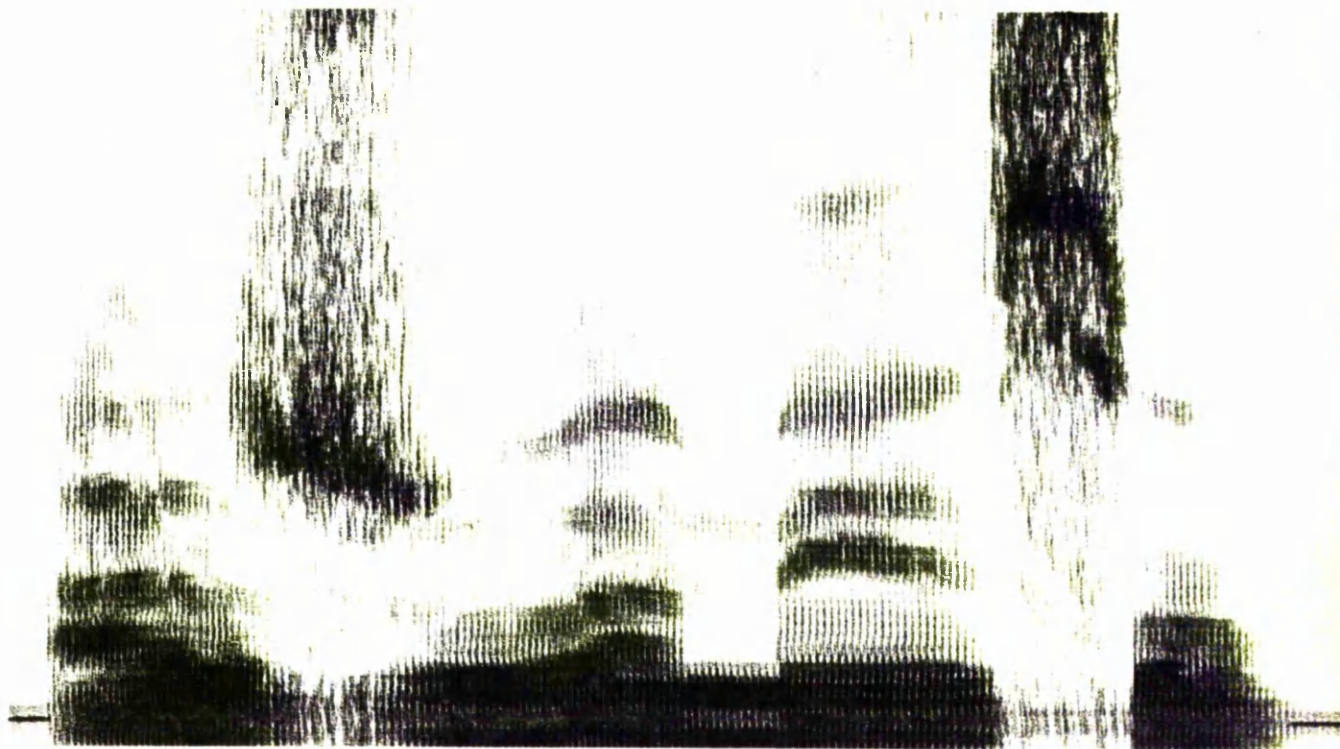
t à m^w á m á s ù s ú

69.



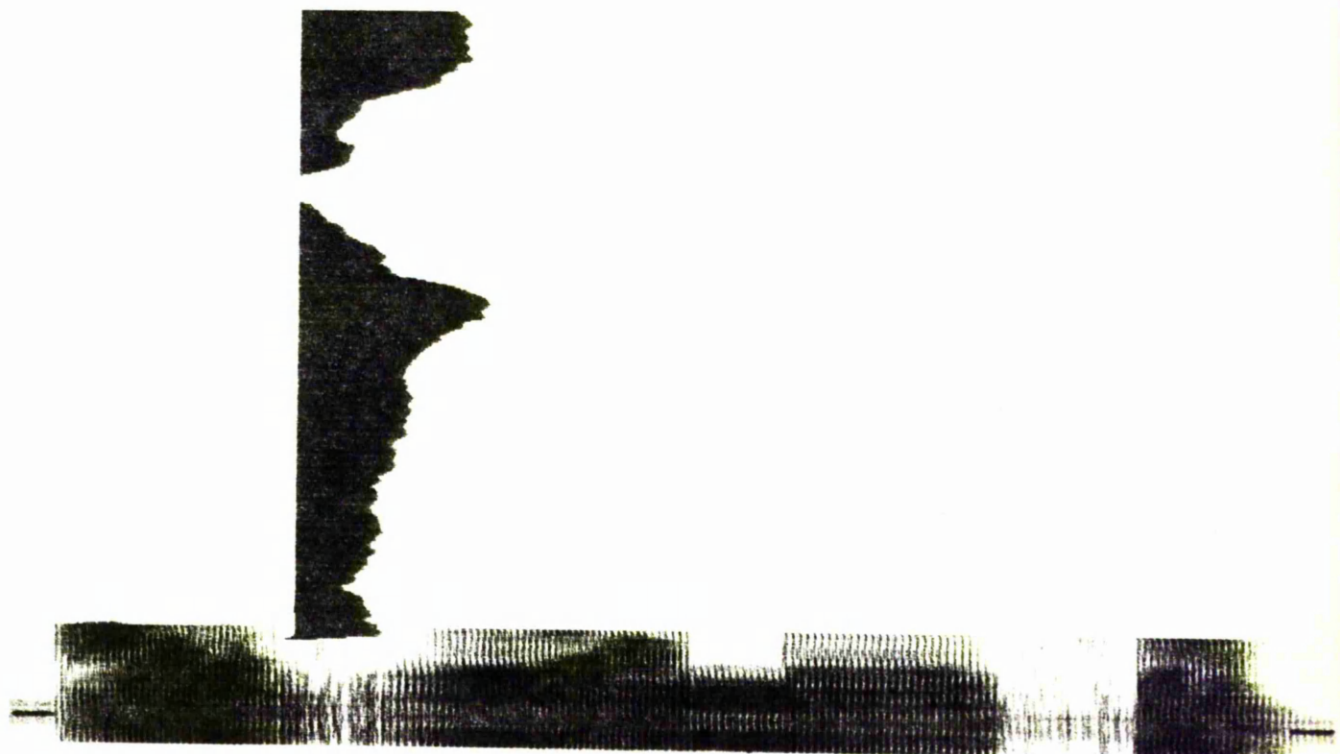
Section of [s]

69.



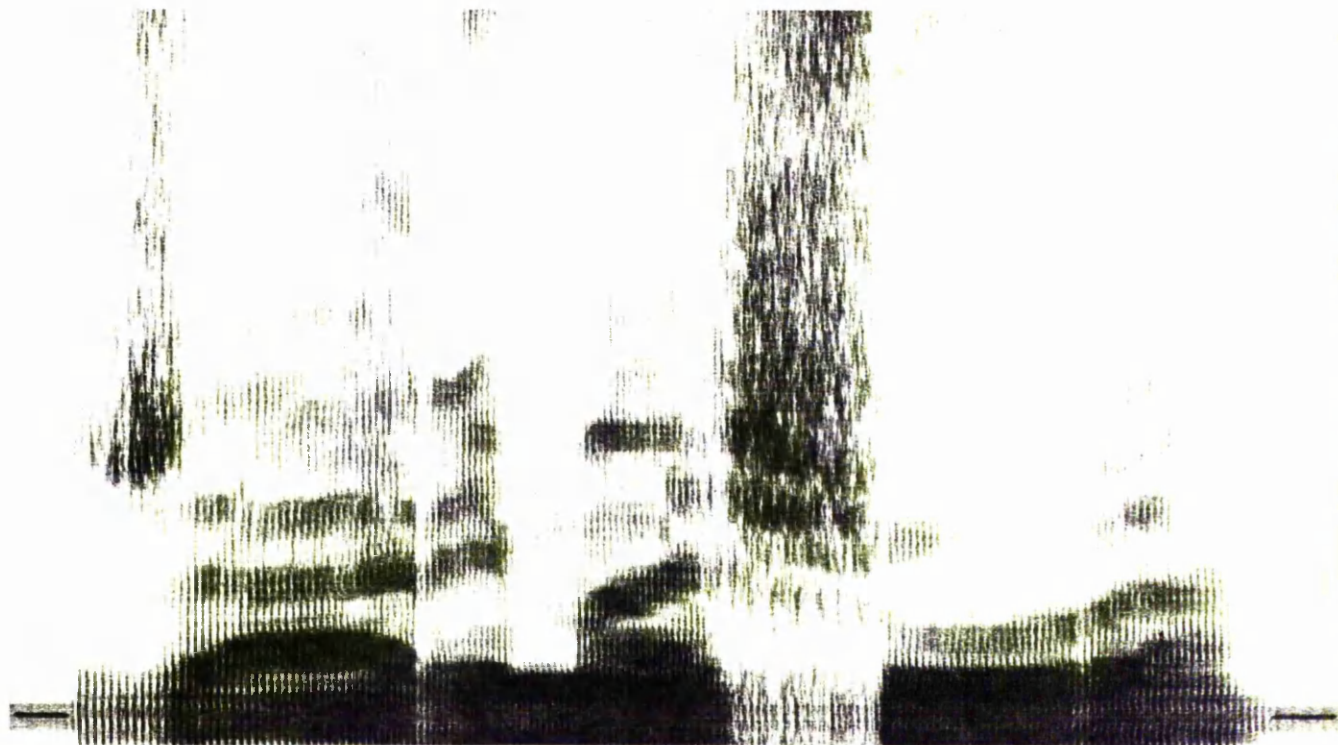
á ẓ ú v à m è s ó

70.



Section of [ẓ]

70.



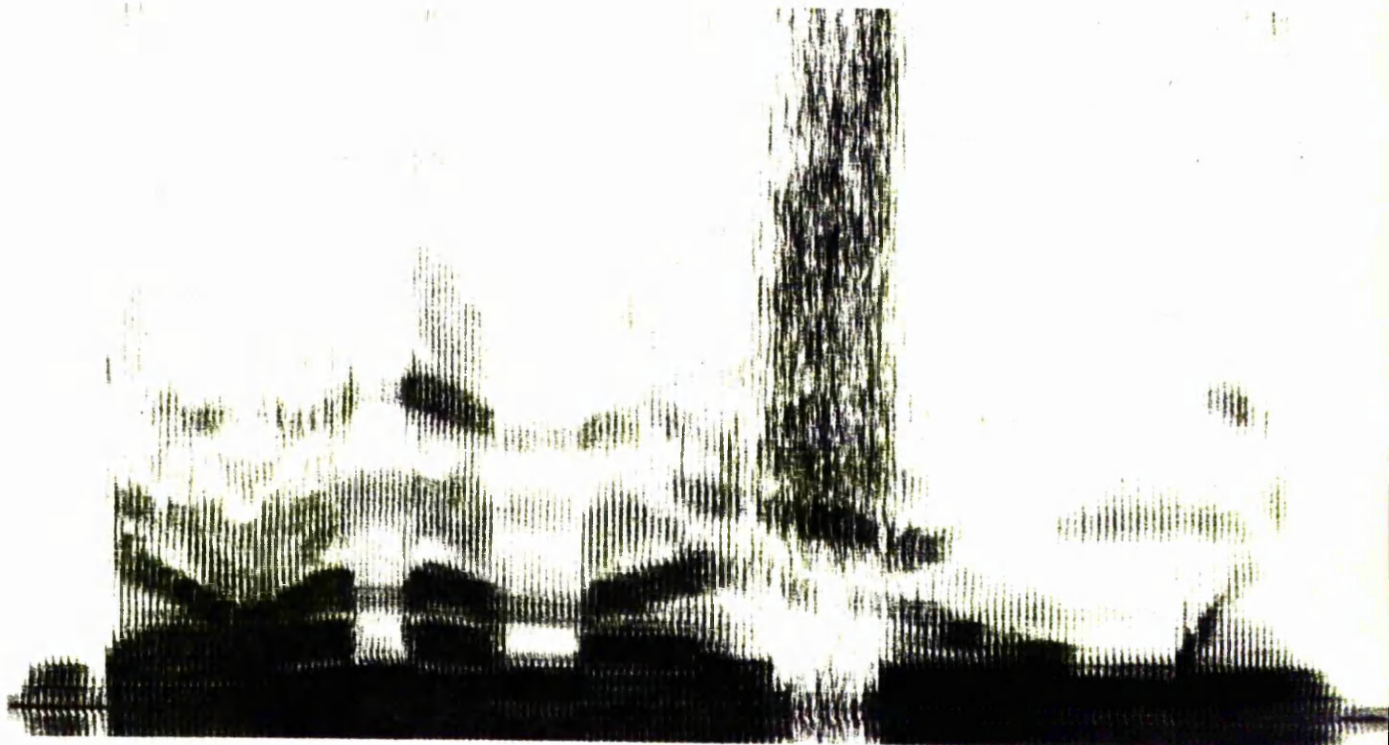
z à á r ì m á j ù r á

71.



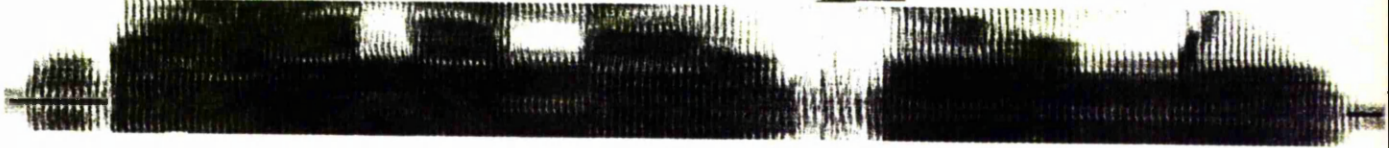
Section of [j]

71.



n d à w á n à m à 3 ù m^w í

72.



Section of [3]

72.

t à x ó r à / n á s i

73.

[x] Section

73.

n d ì n ó đ á ý ó w à

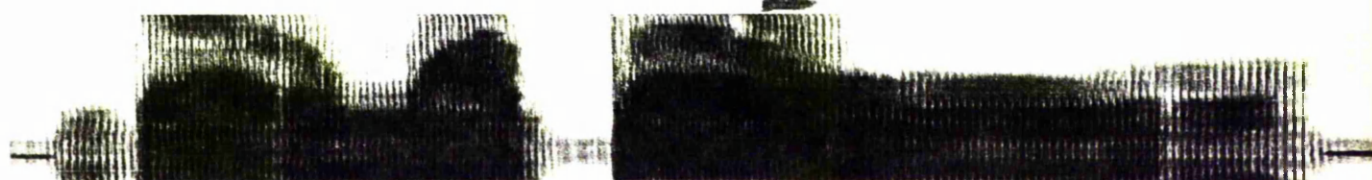
74.

Section of [ý]
74.



n d à r ú m^w a' p á ã ù r ò

75.



Section of [ã]

75.



m à pf ú m ó á w á n ã à



u

k

u

h

77.

a

r

u

by

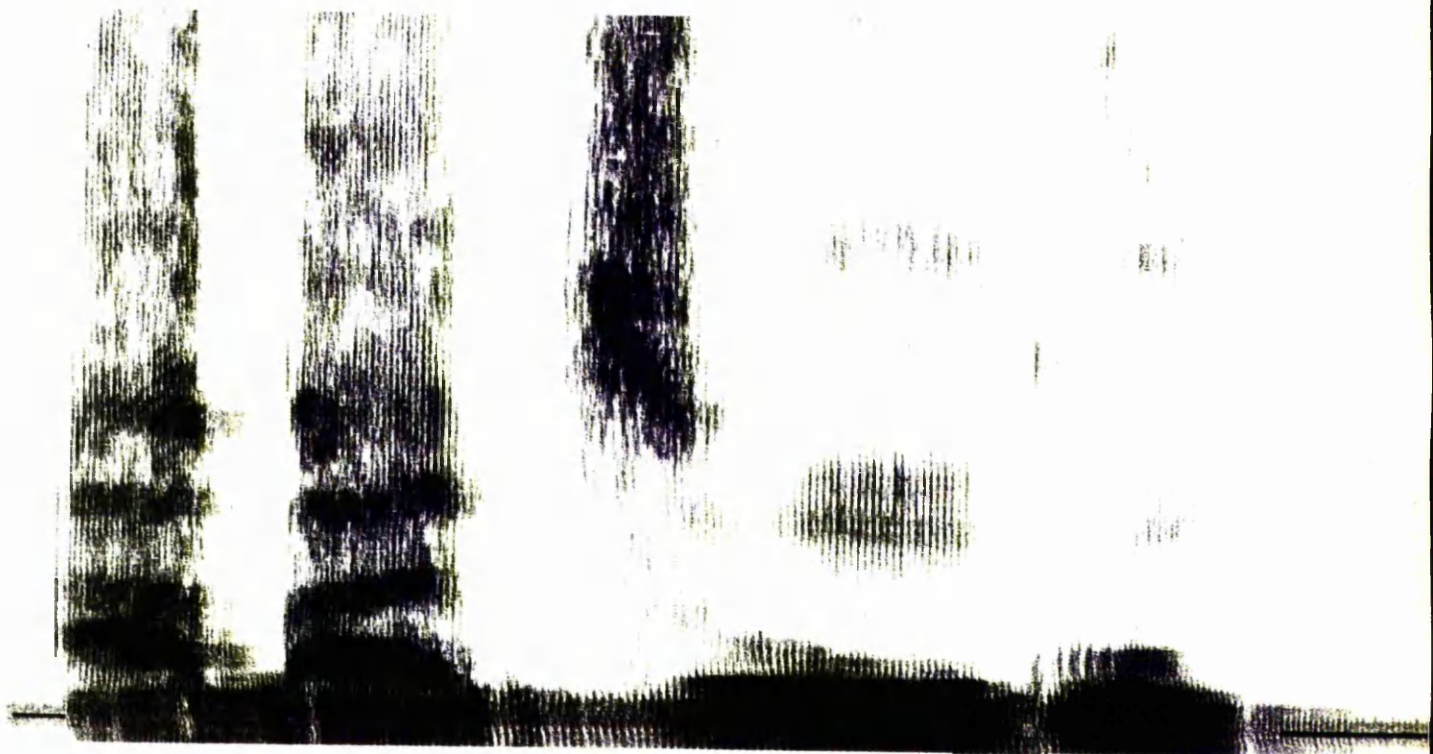
a

m



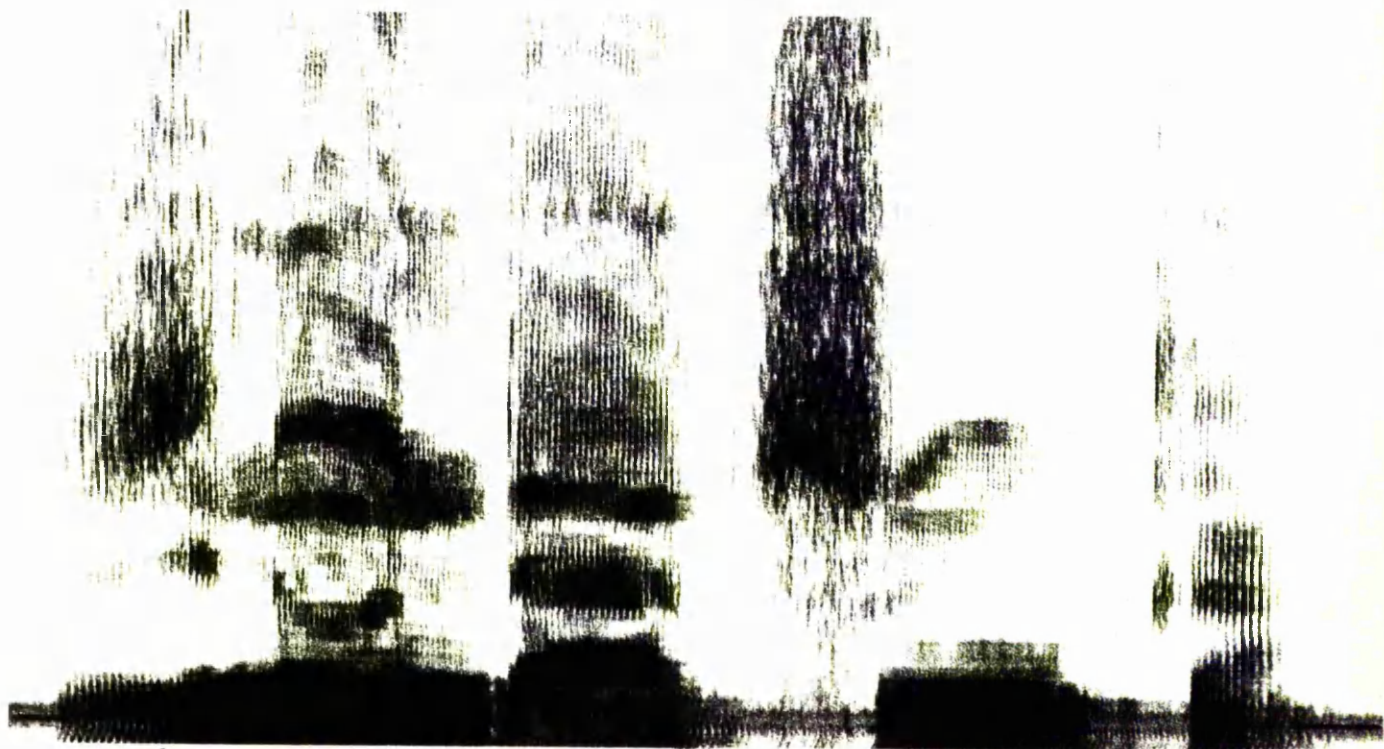
ndĩ n ó ɗ à 'ú ŋ g à

ts
78.



á 6 à dz ù ŋ g ú

79.



z í n ò nd à ts ú k à

80.

ú

k

ù

đz

81.

á

đ

ó

n

đ

i

n



n d à m ú p à t f ú m í



ú

r

ù

đ

83.

á

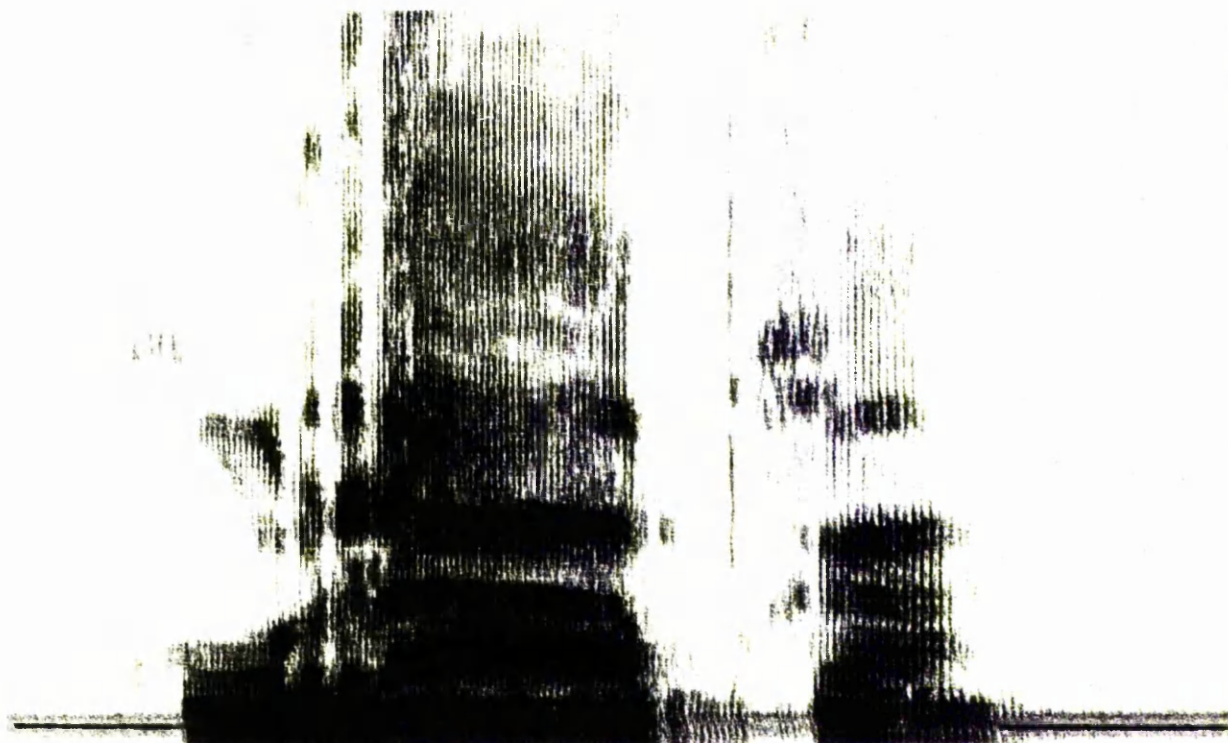
m

á

h

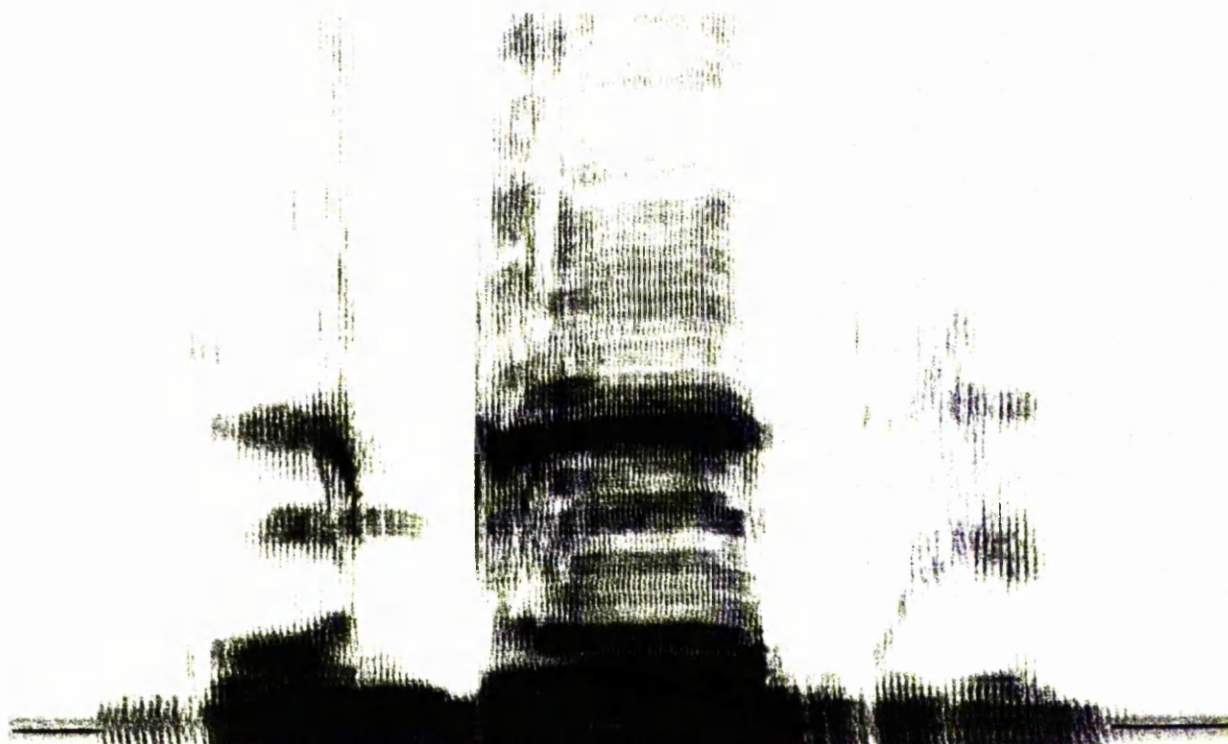
à

nđ



k ù r à p x à

84.



g ò n d ó b ỹ è

85.



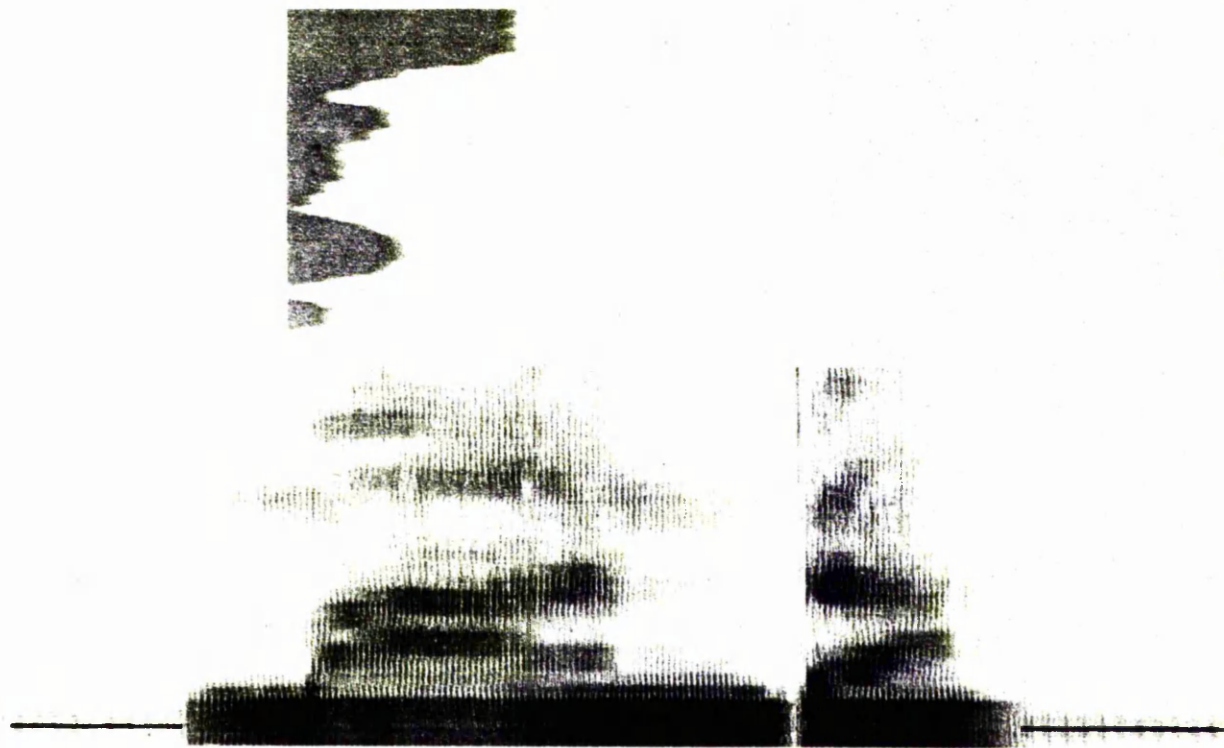
tj

ú

m

í

86.



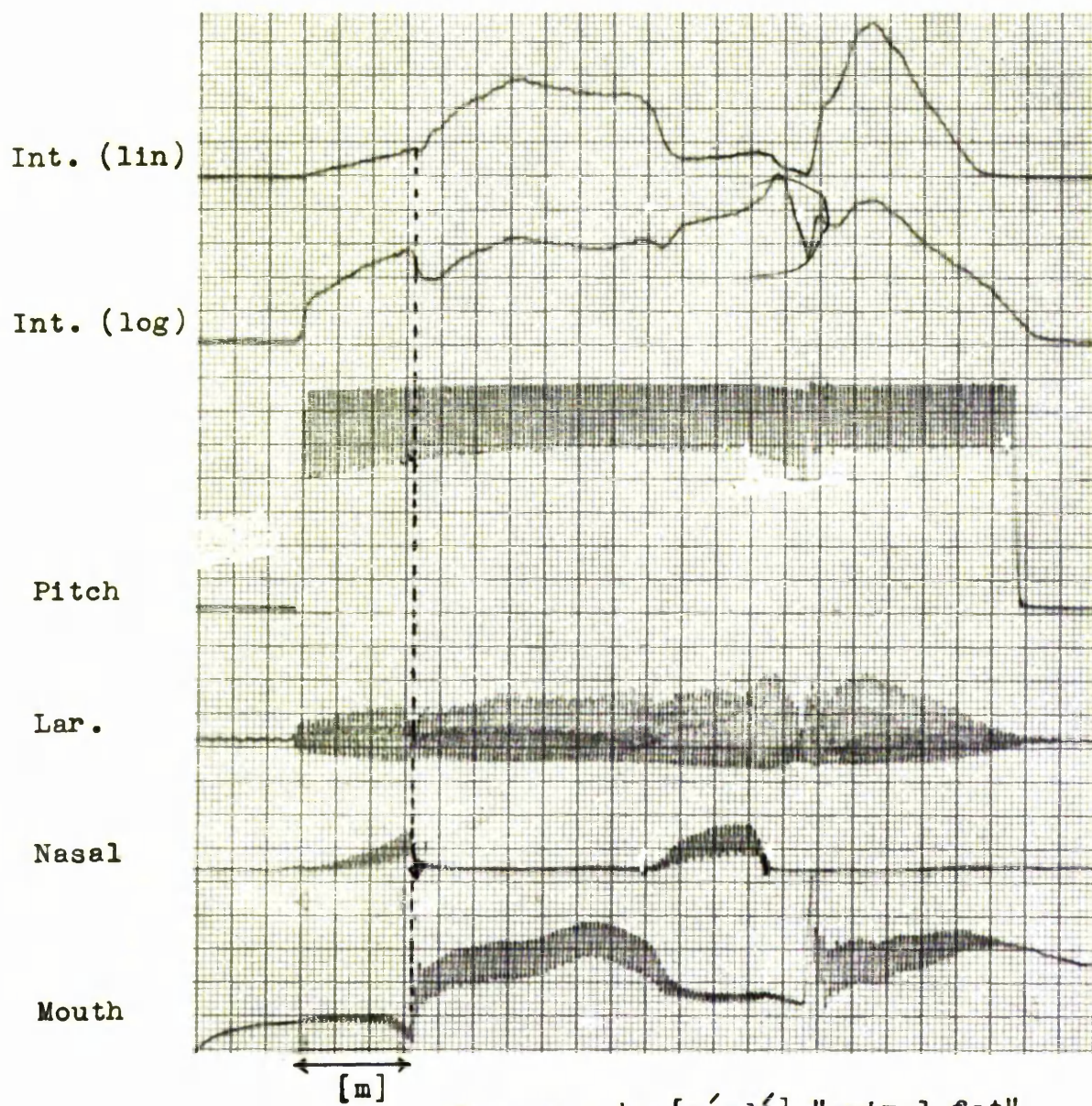
m á n d á

87.

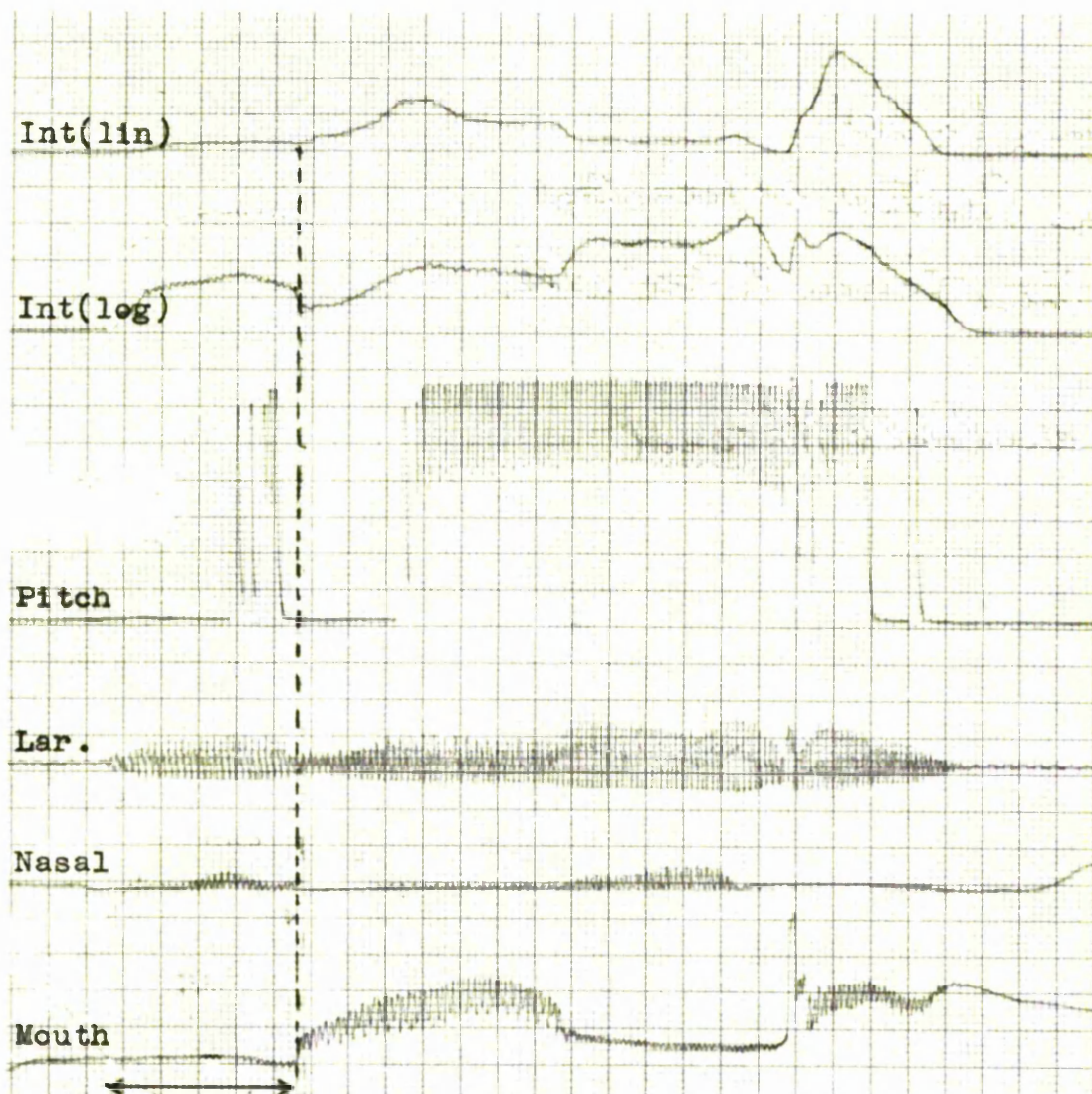


m á n d á

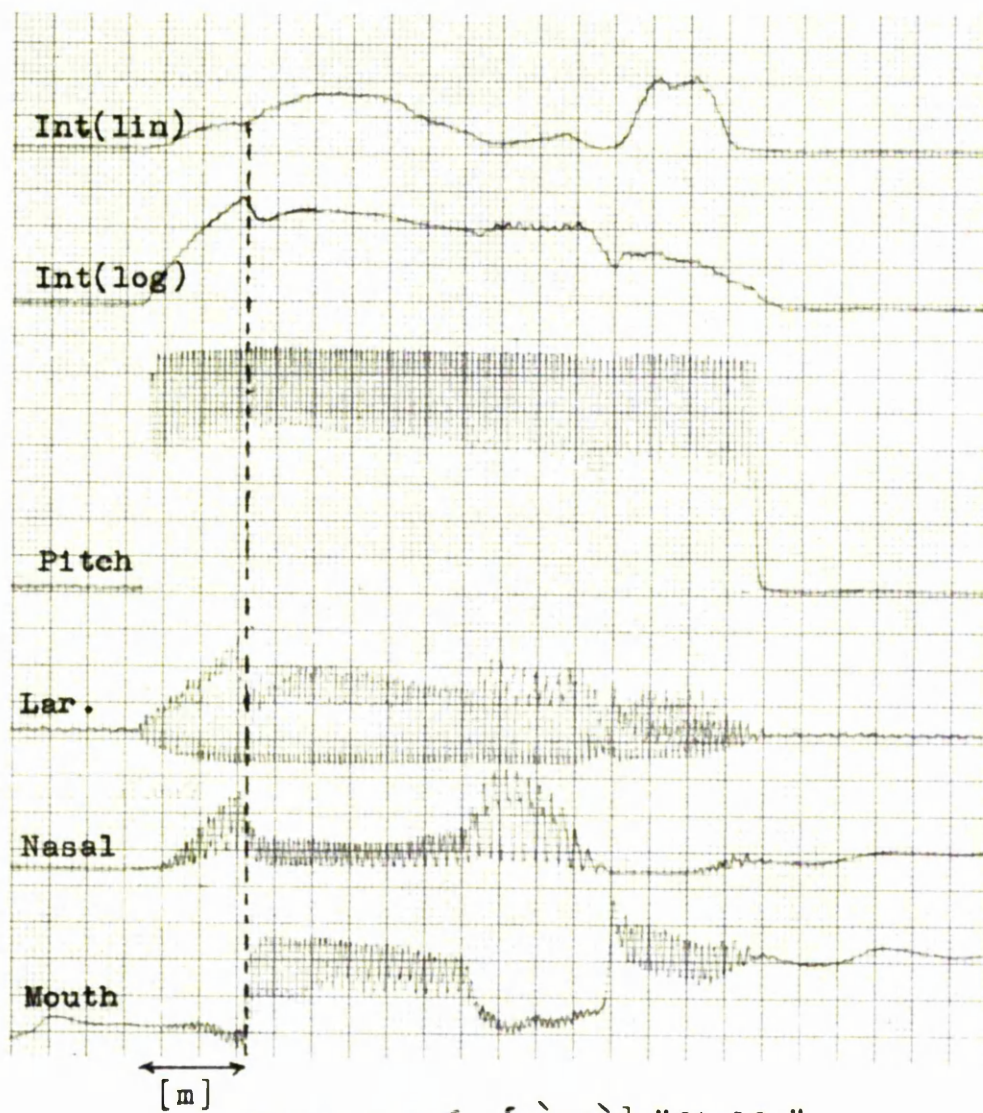
88.



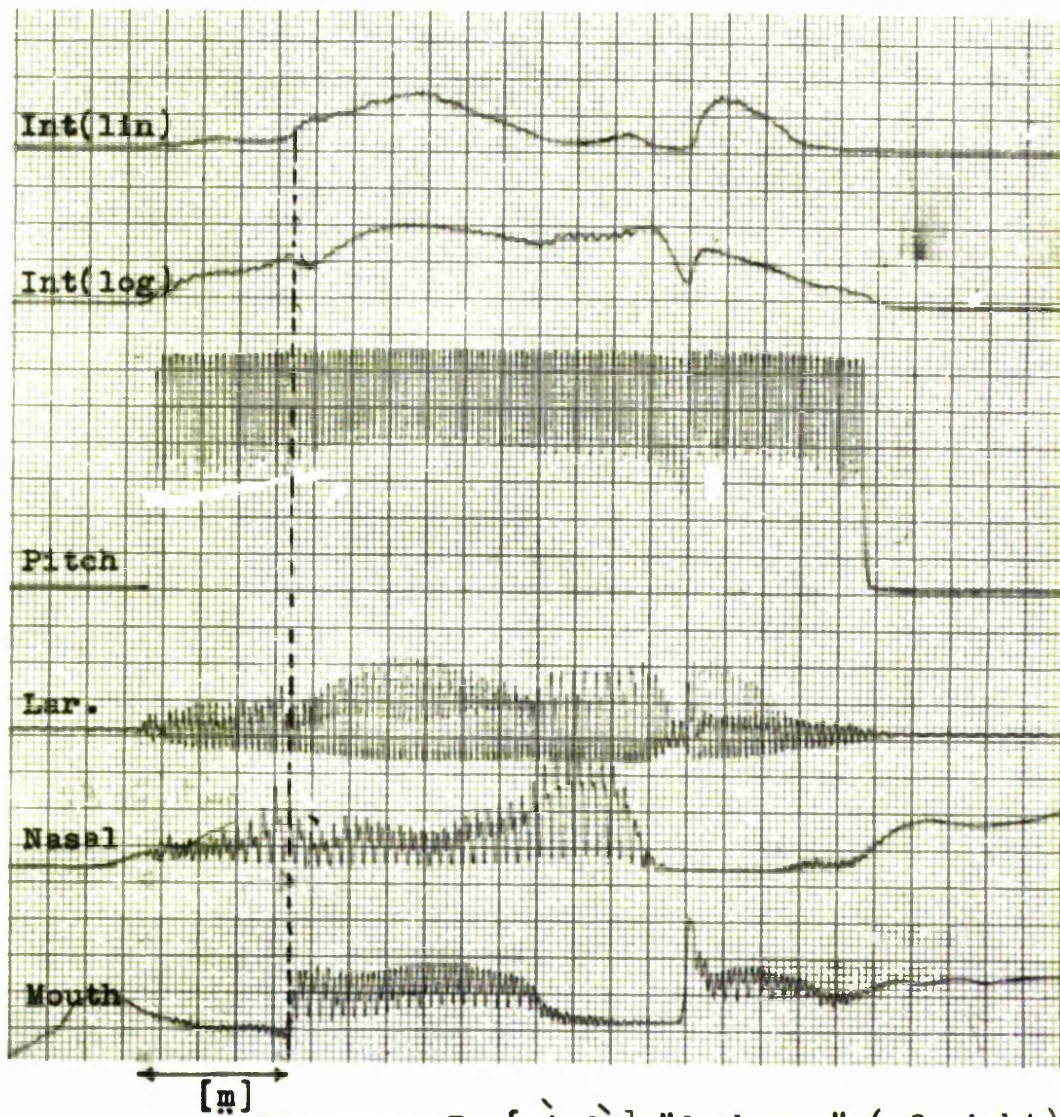
Mingogram 4. [mándá] "animal fat"



[m] Mingogram 5. [mándá] "fork in branch of tree"



Mingogram 6. [mìndà] "fields"



Mingogram 7. $[m̩indò]$ "darkness" (of night)



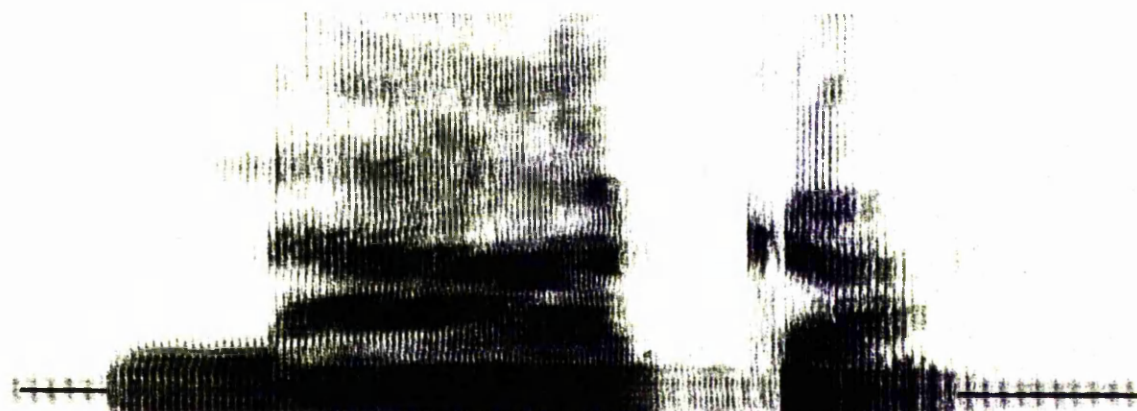
n

`a

k

`a

89.



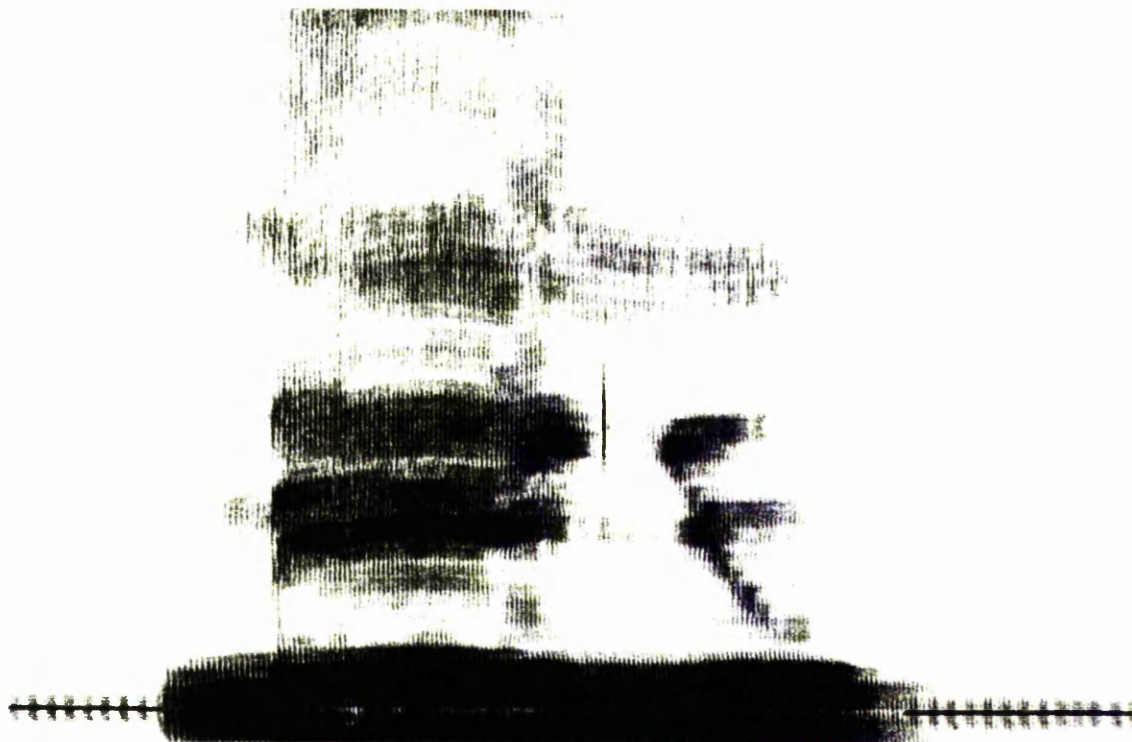
n

`a

k

`a

90.



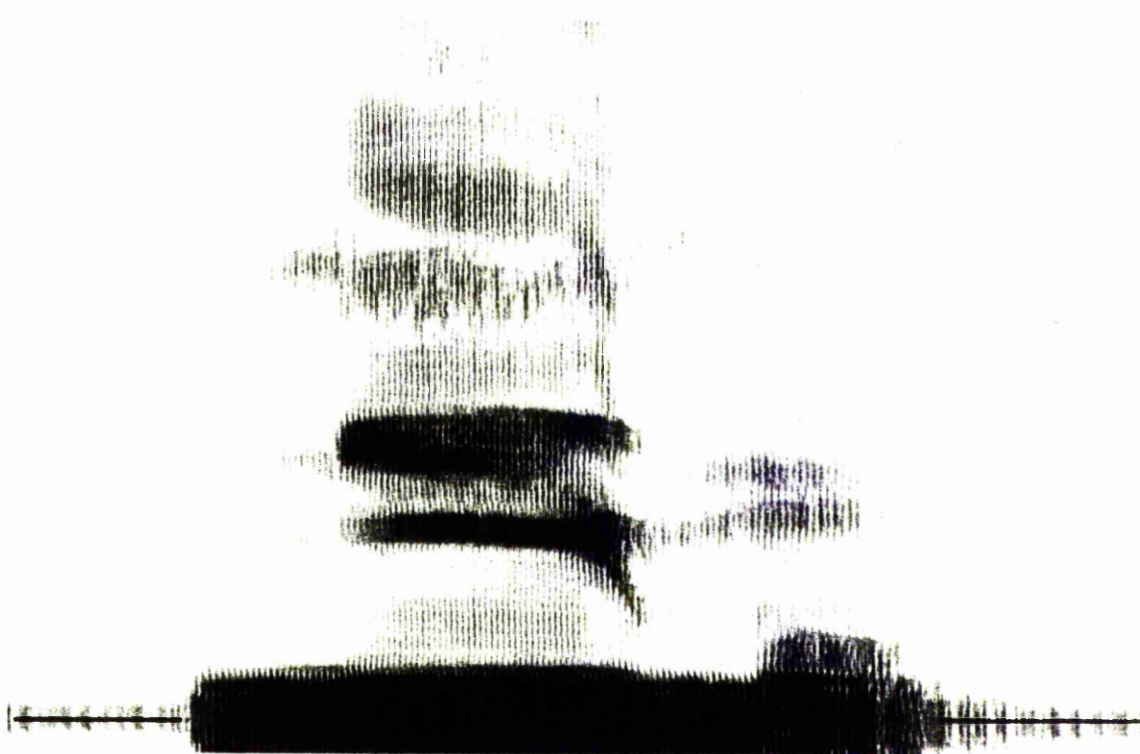
m é n ú

91.



m è n ú

92.



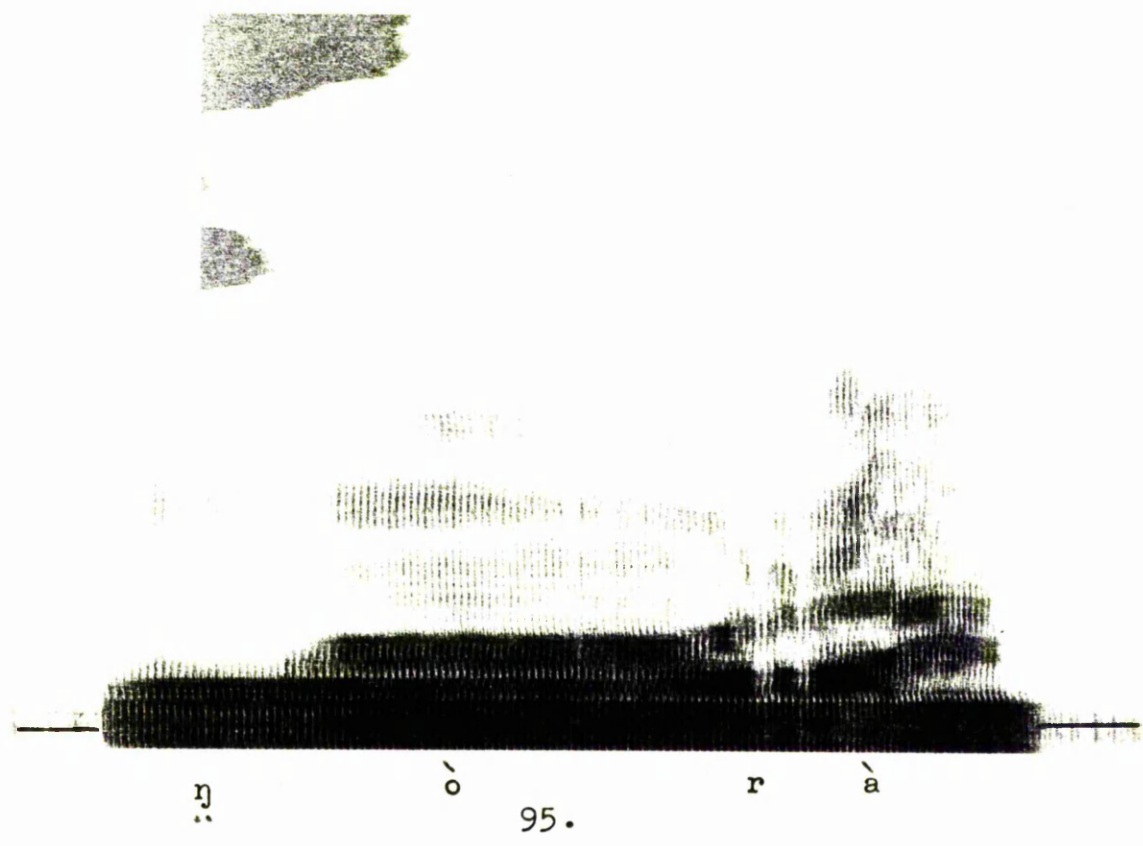
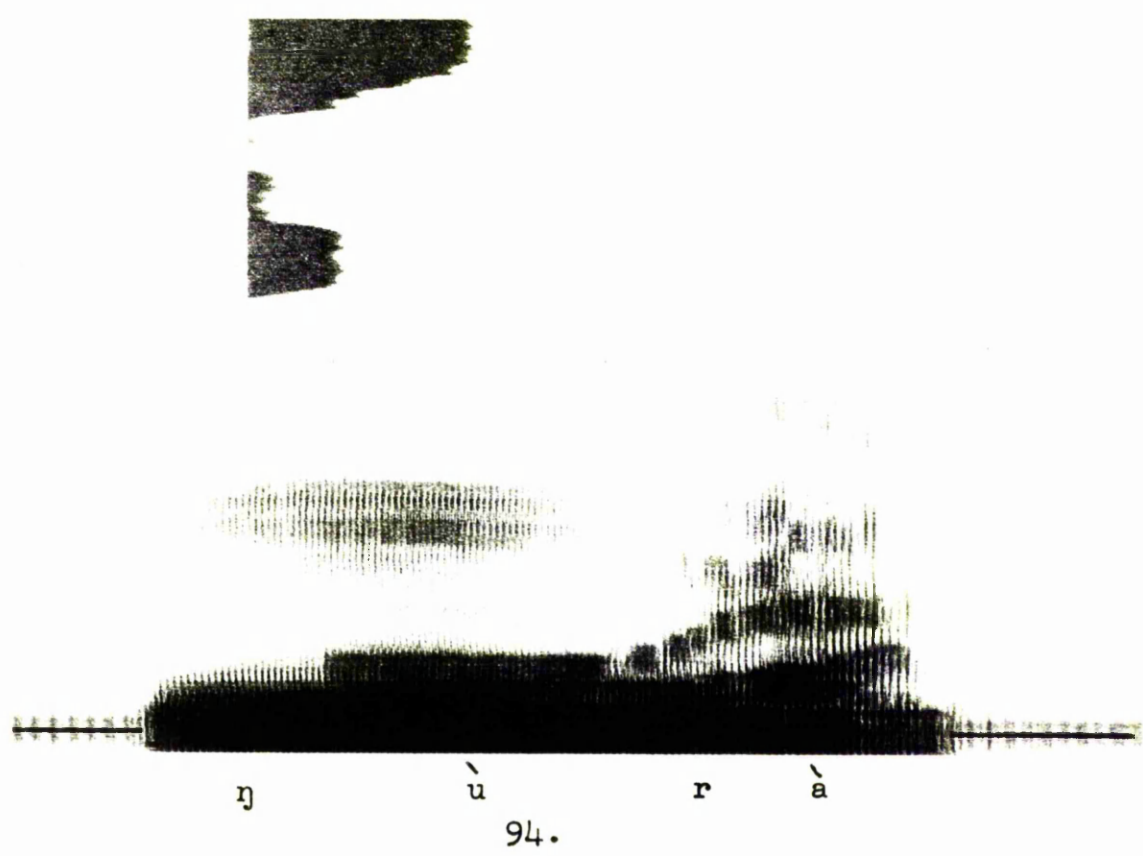
p

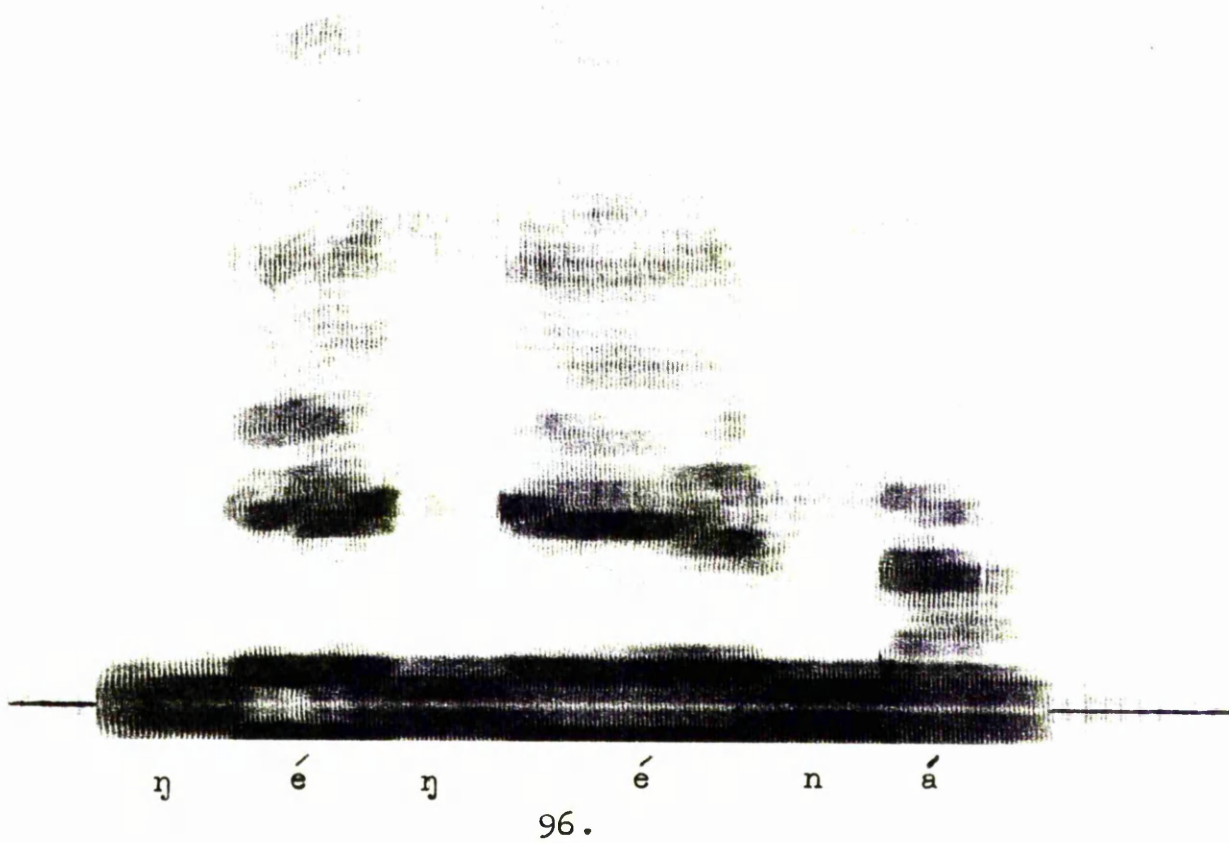
i

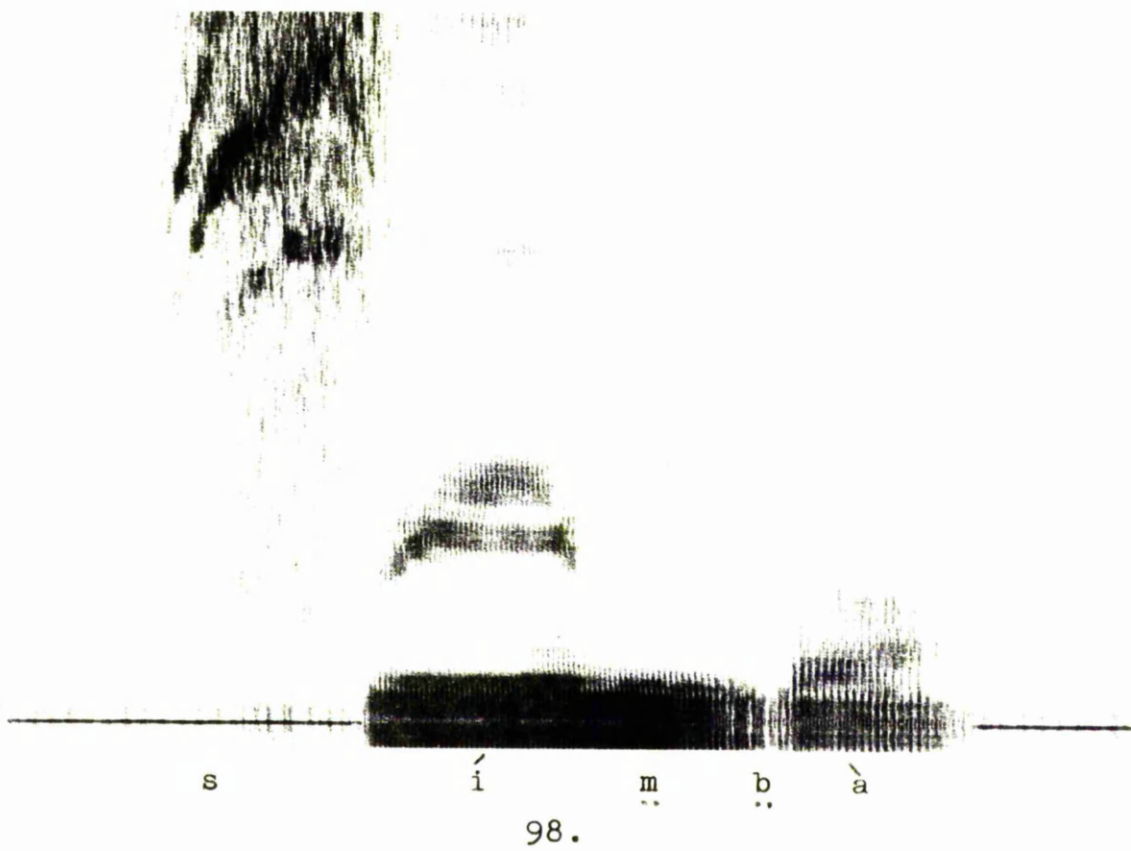
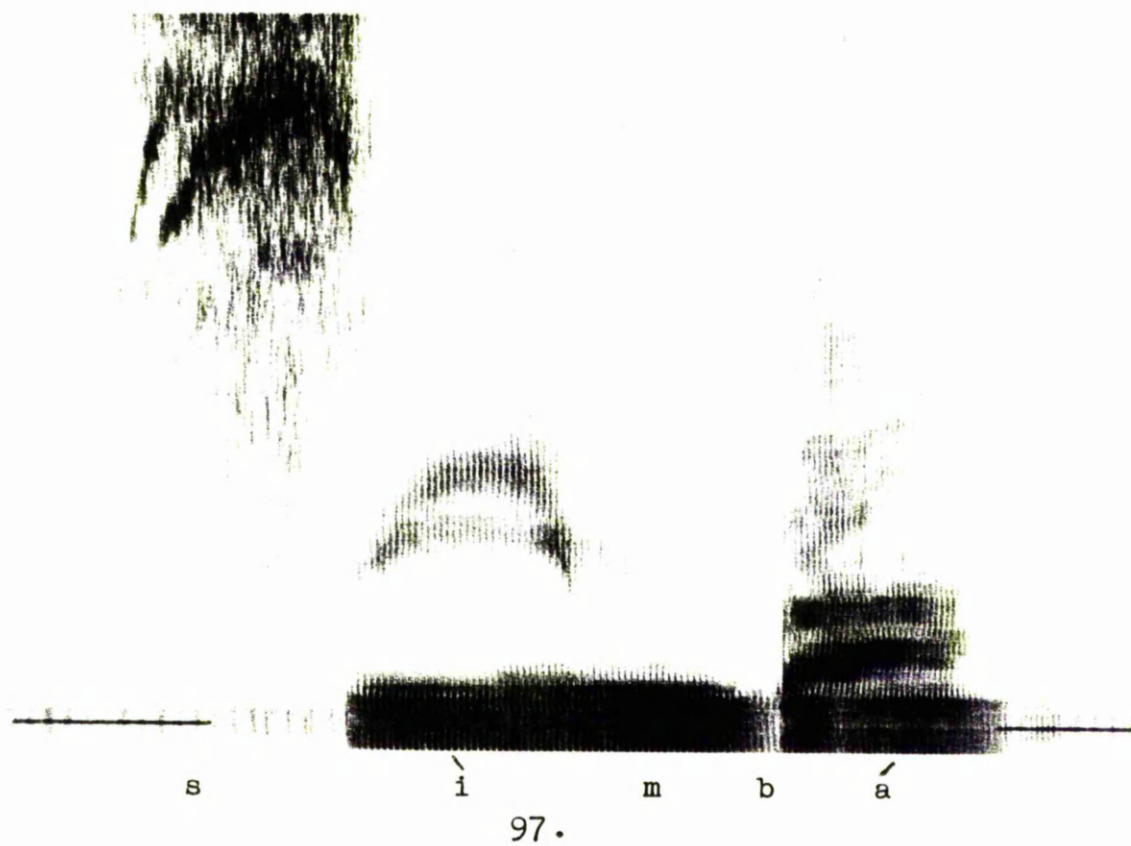
93.

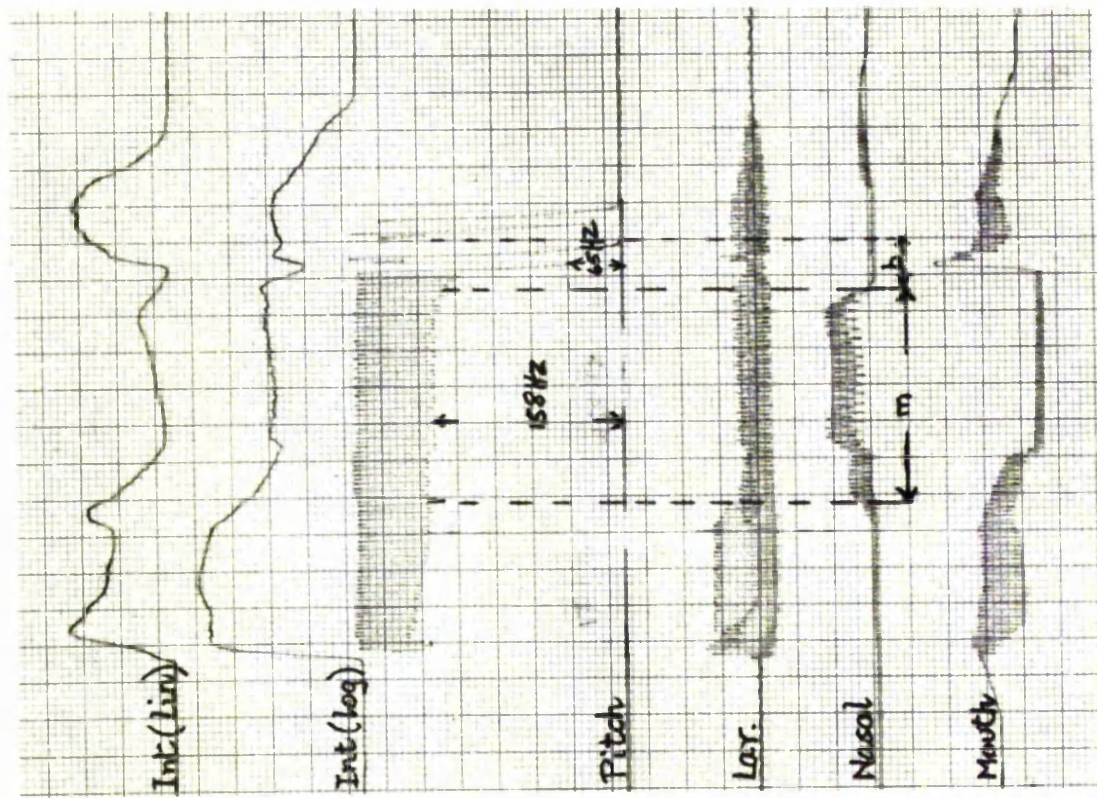
m

o

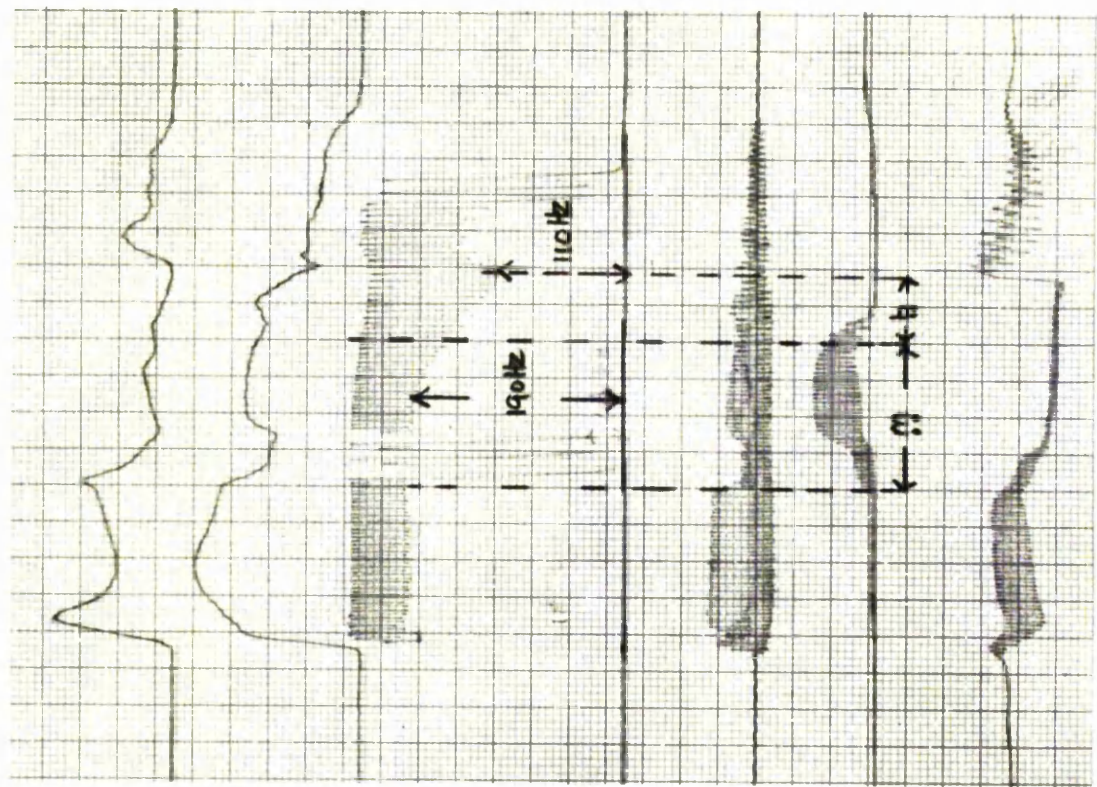




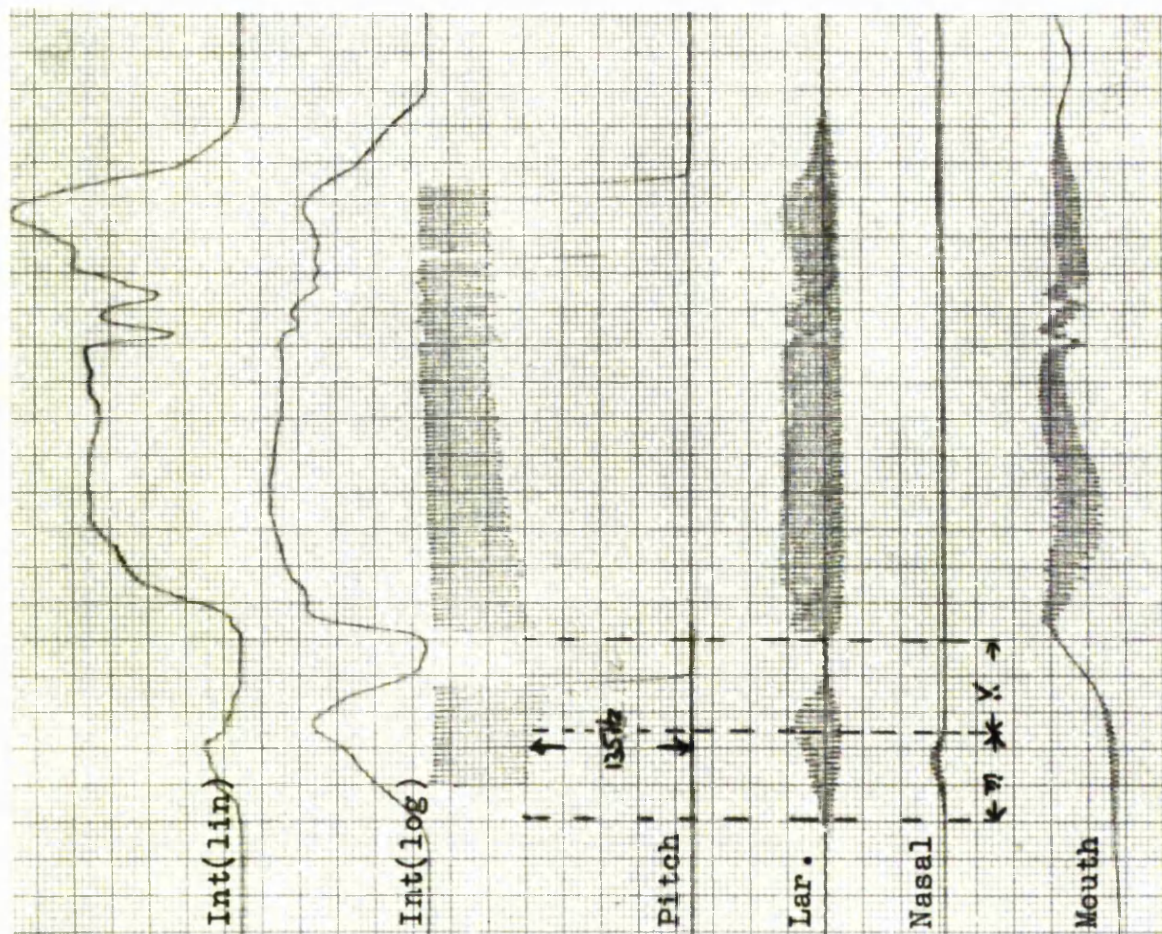




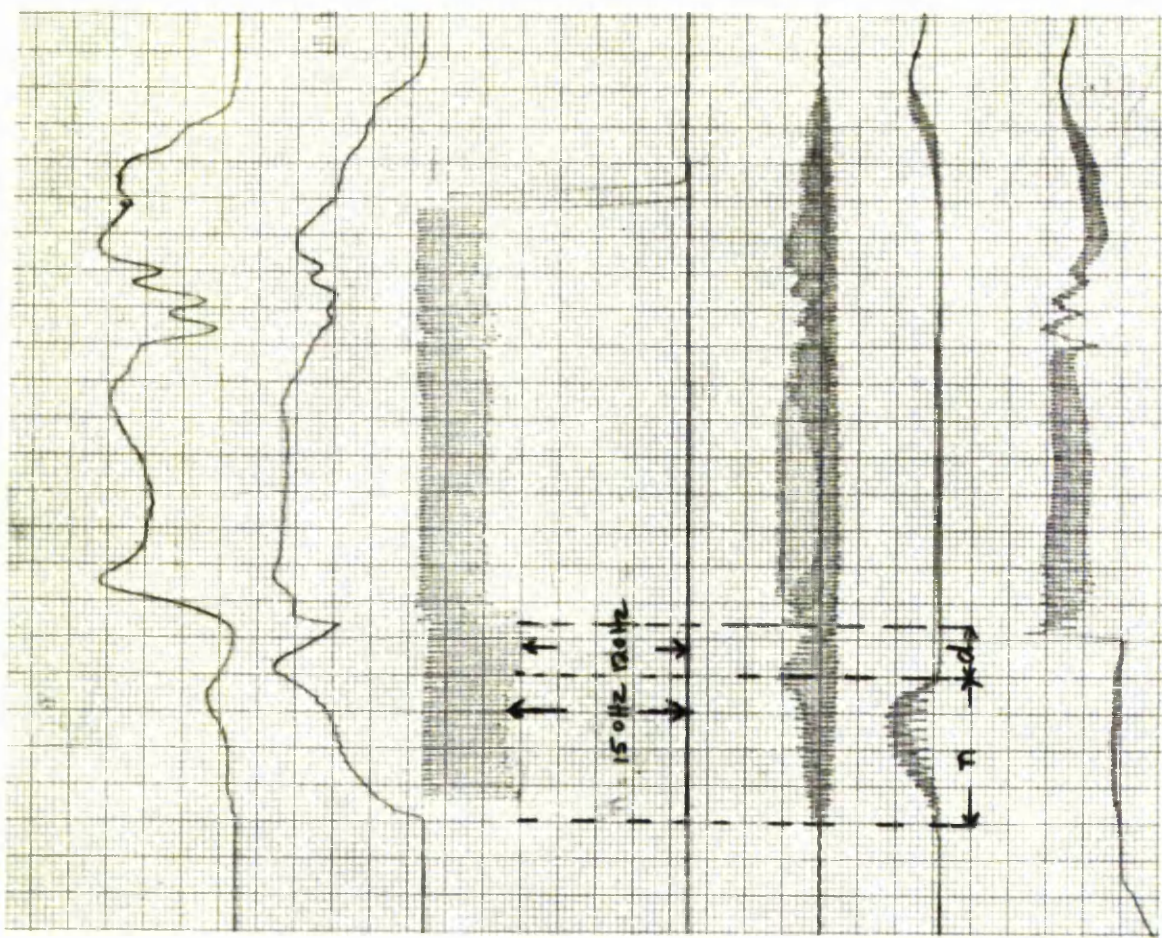
Mingogram 8. [simba] "strength/power"



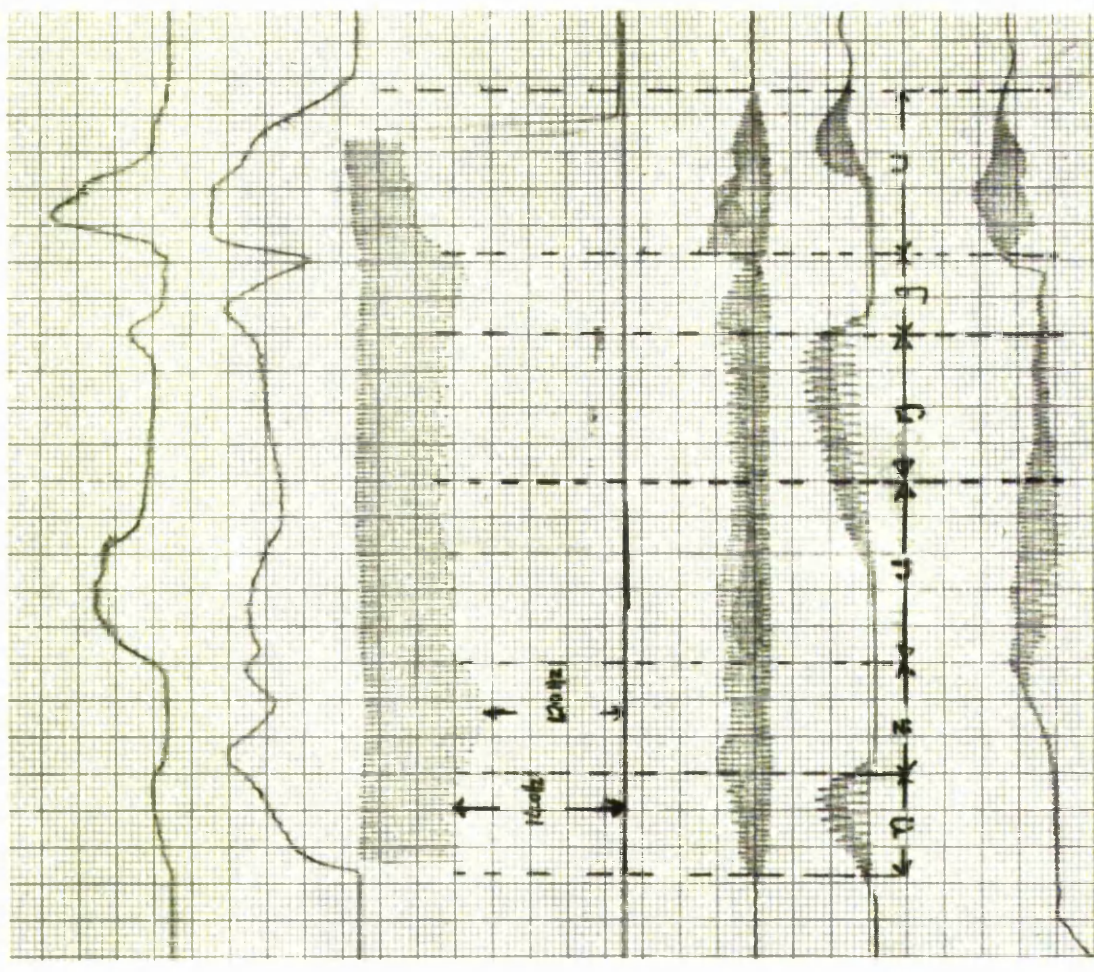
Mingogram 9. [simba] "genet"



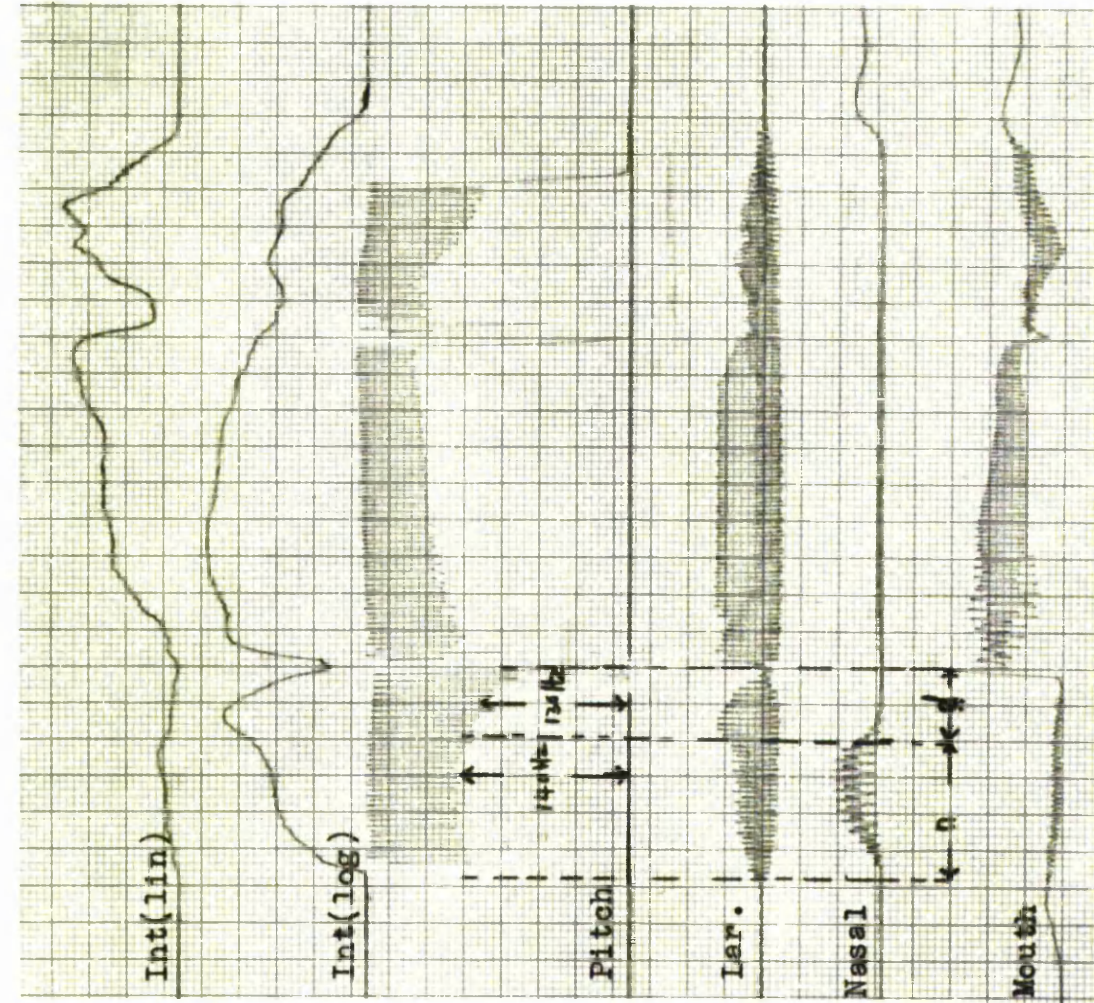
Mingogram 10. [mivura] "water/rain"



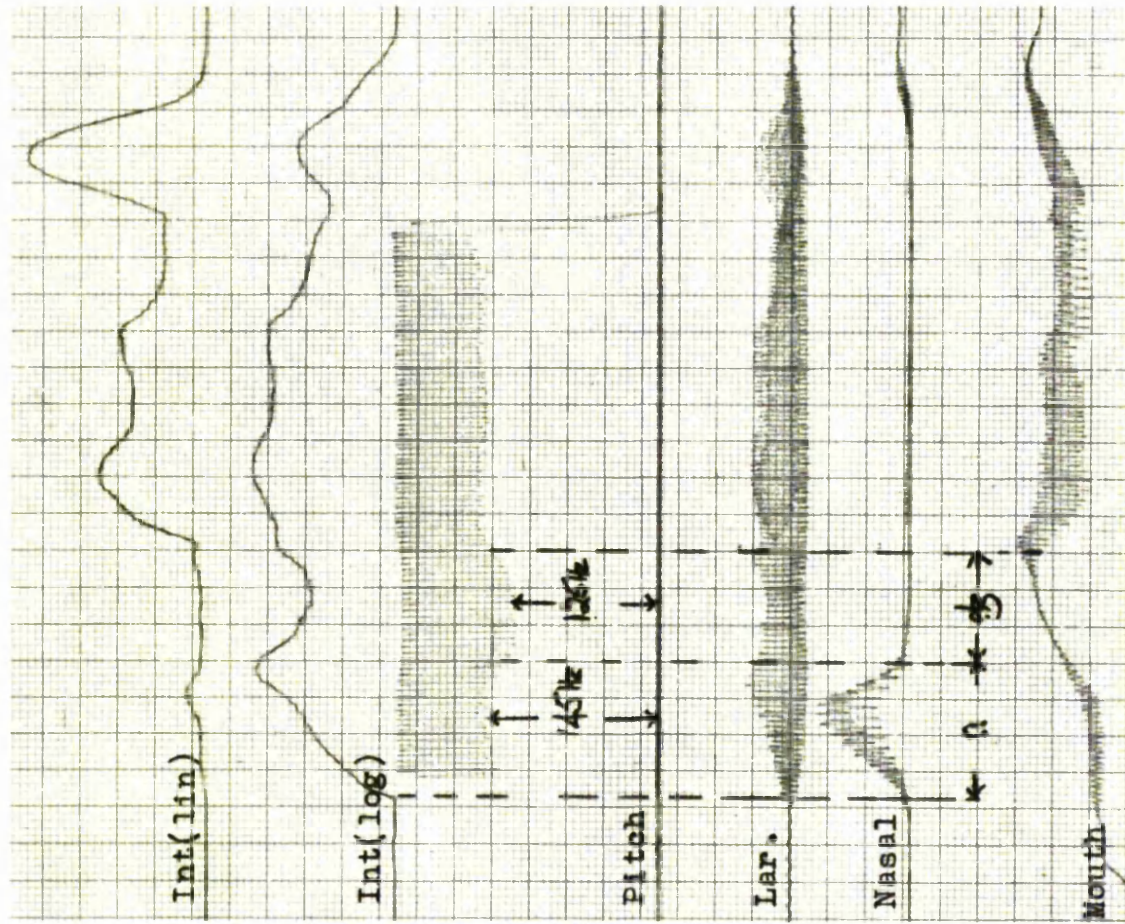
Mingogram 11. [ndiro] "plate" (n)



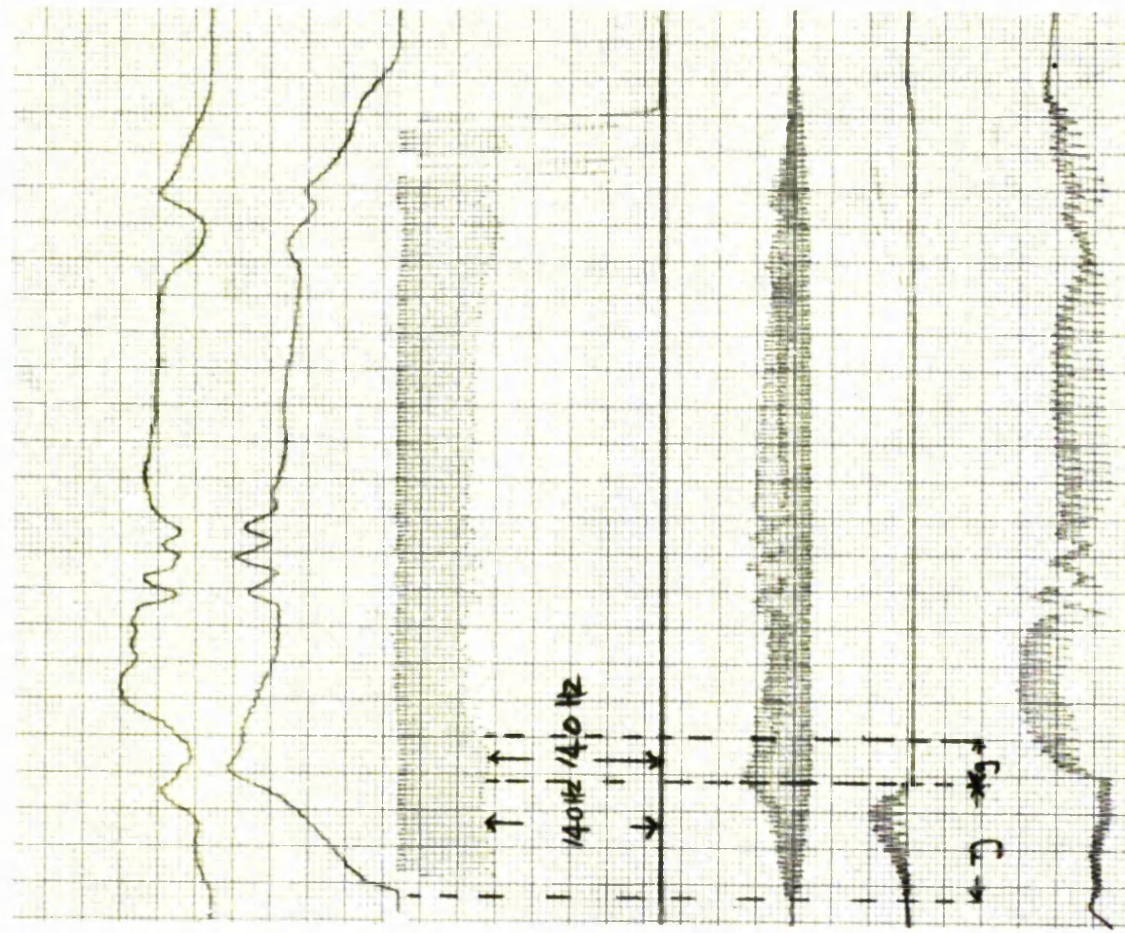
Mingogram 13. [nɛ̀nɛ́gú] "peanuts"



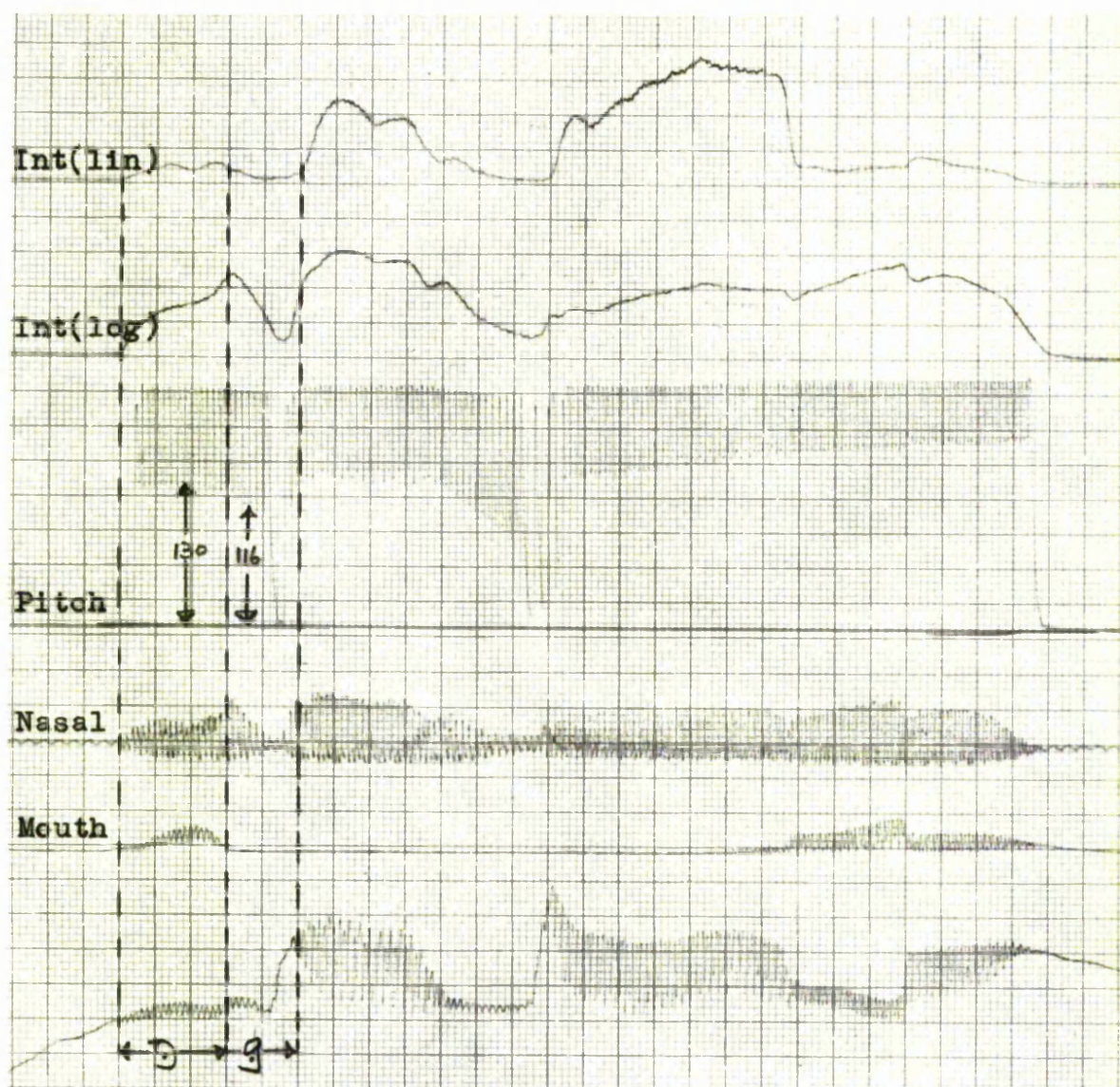
Mingogram 12. [ndíró] "it is the one"



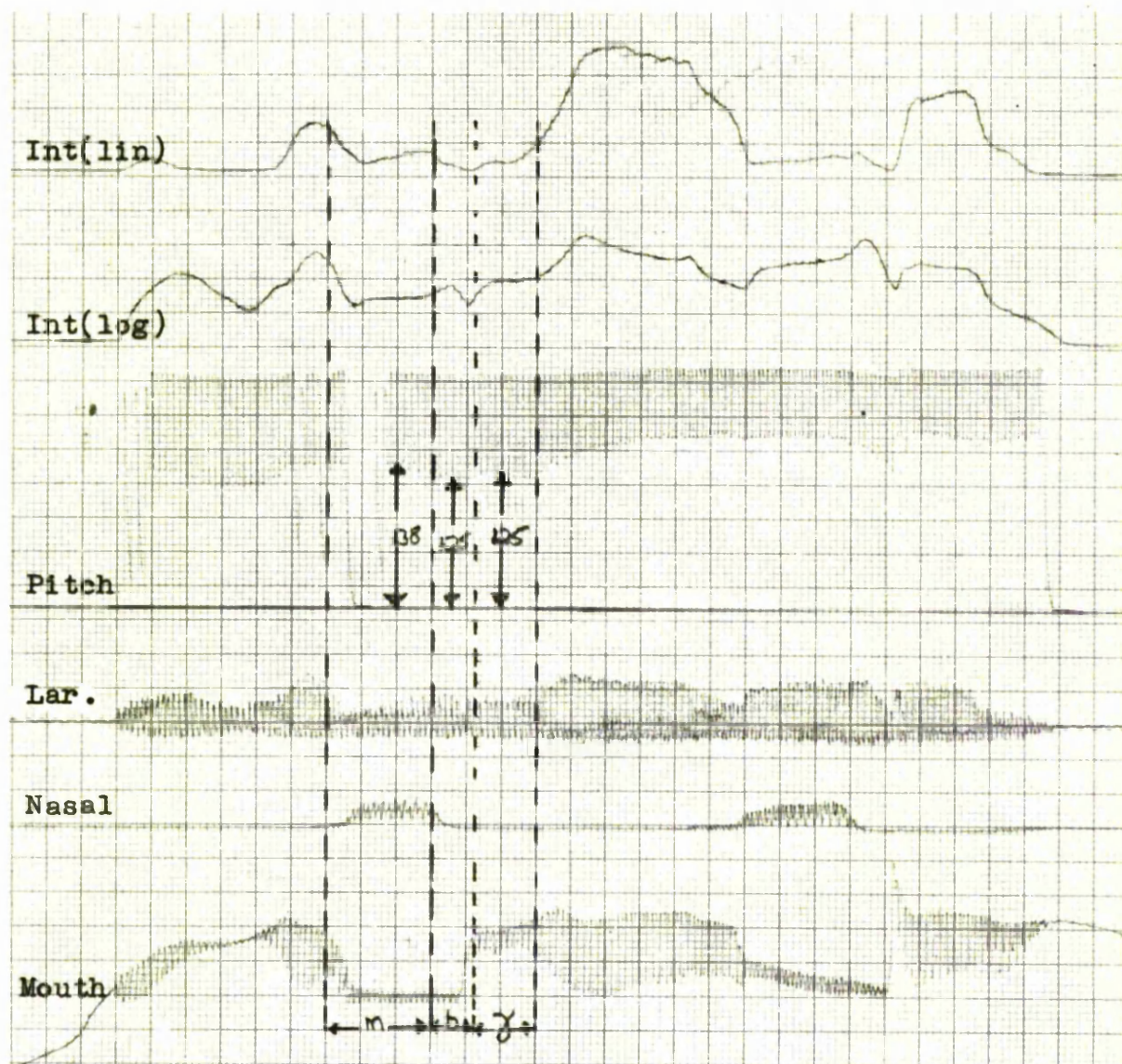
Mingogram 16. [ndʒivá] "dove"



Mingogram 17. [ɲgúrùvè] "pig"



Mingogram 18. [ɲgùbání] "child's intestinal disorder"



Mingogram 19. [zĩmbɔréndé] "coward"

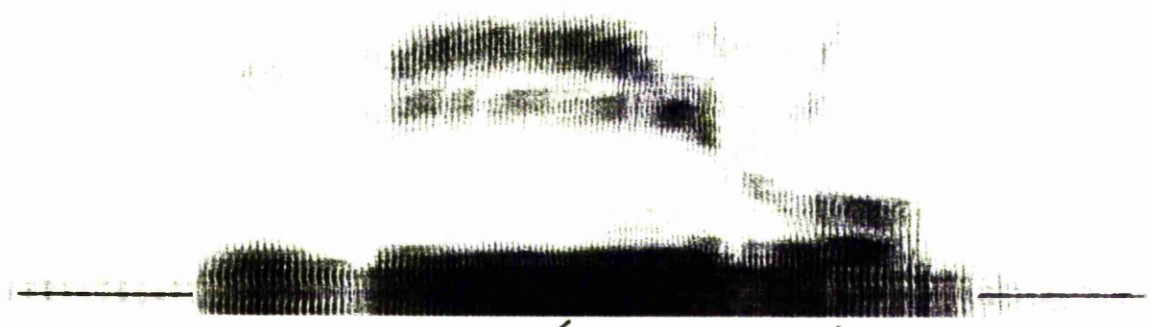


ndi n o d a m v u r a



n d í r ó

100.

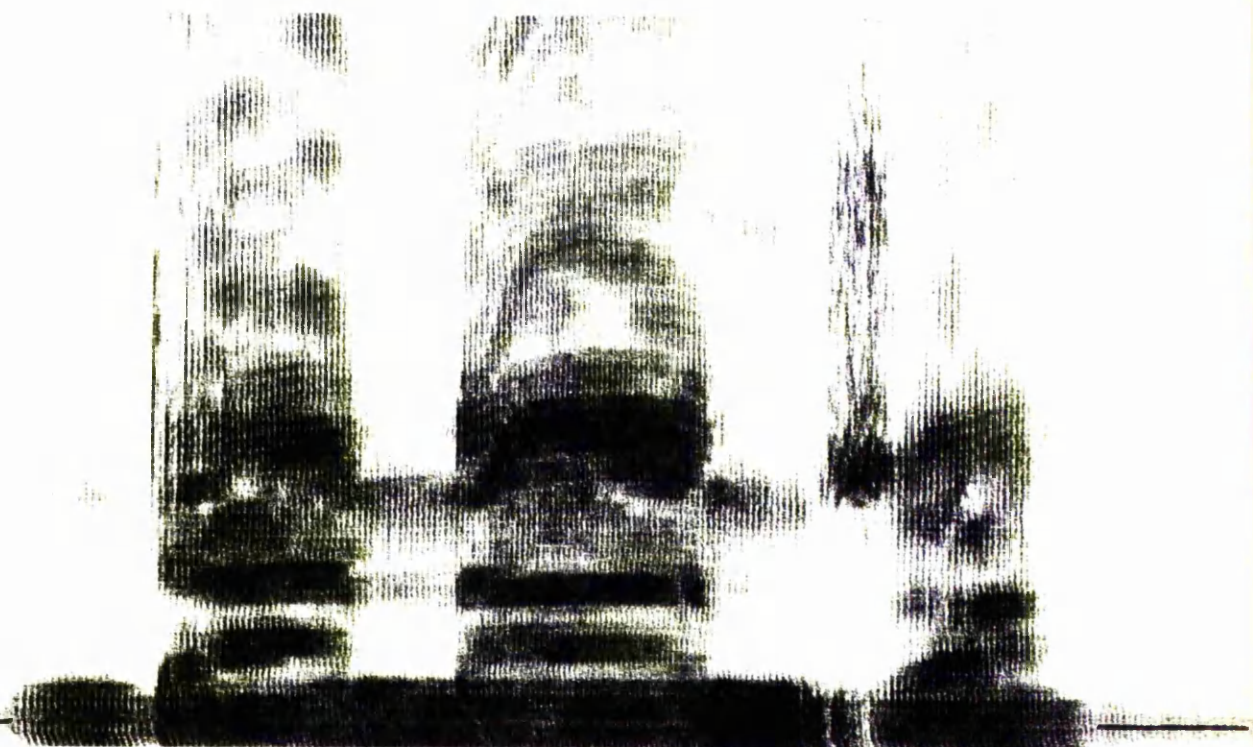


n d í r ò

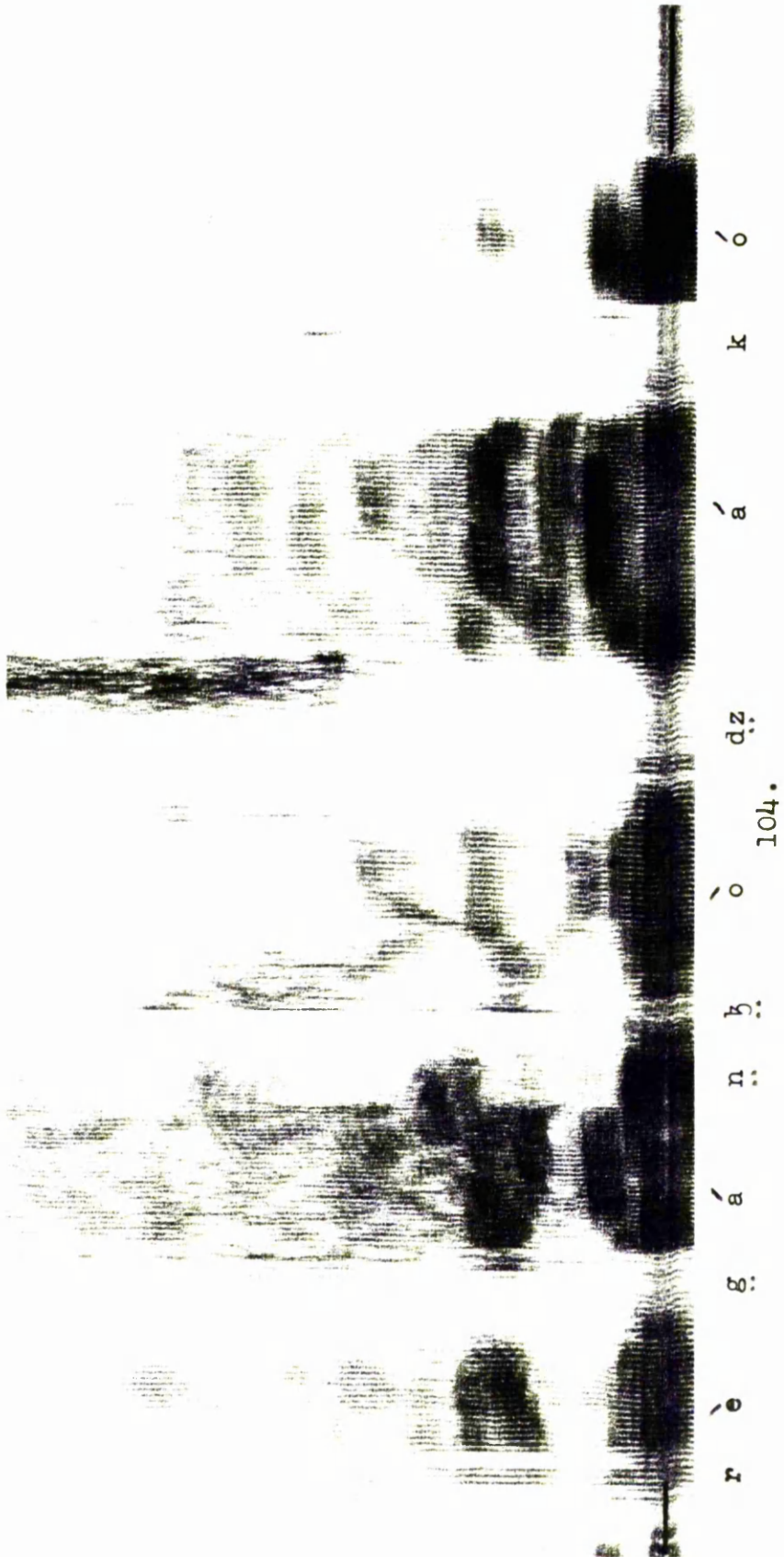
101.

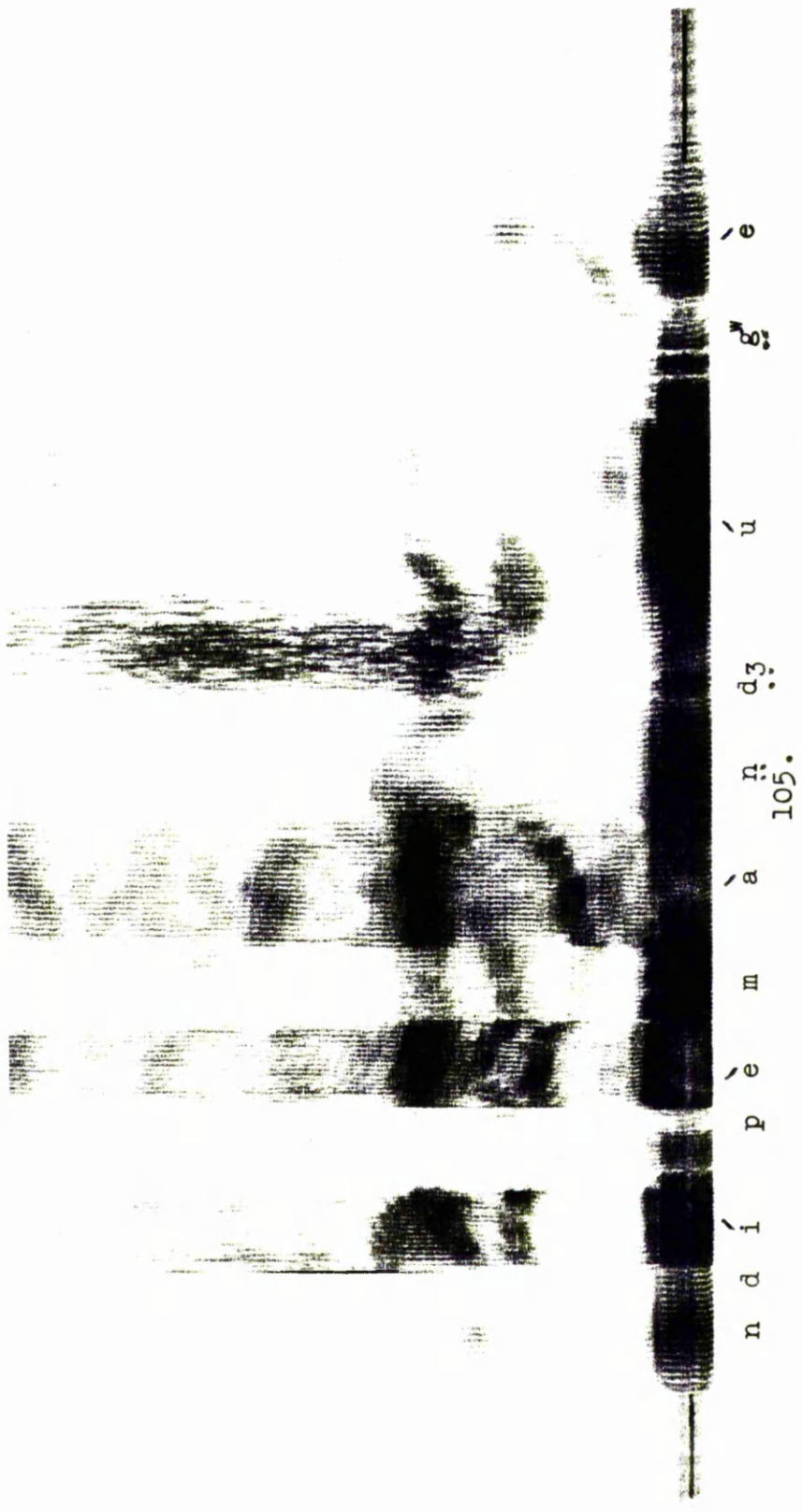


k è t á n z ú m à
102.



n d à n á n z á
103.





,

g

u

dj

n
105.

a

m

e

p

i

d

n



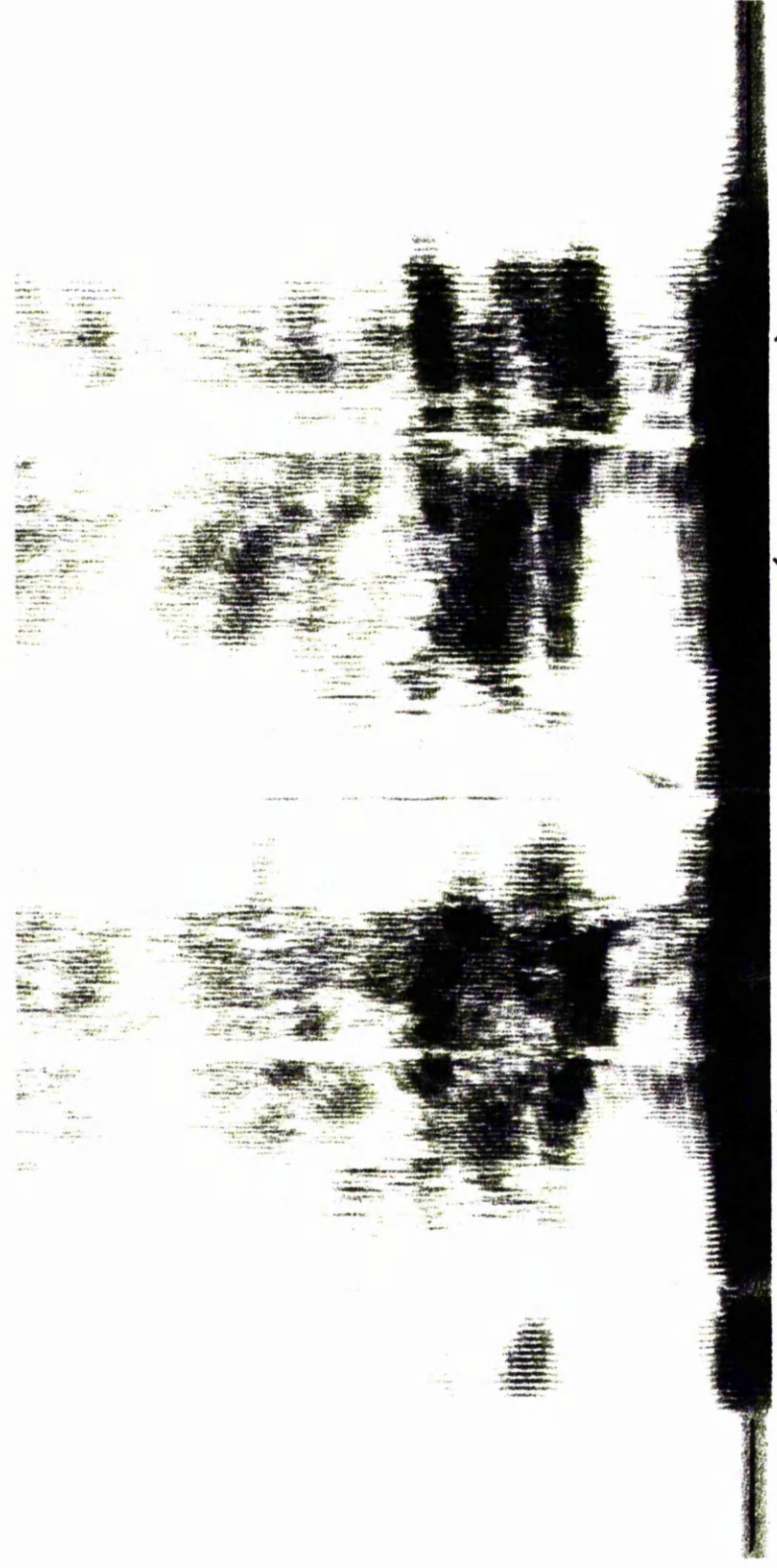
η g ù ù r ù v è

106.



η g ù b á n í

107.



e

r

i

y

m. b. 108.

e

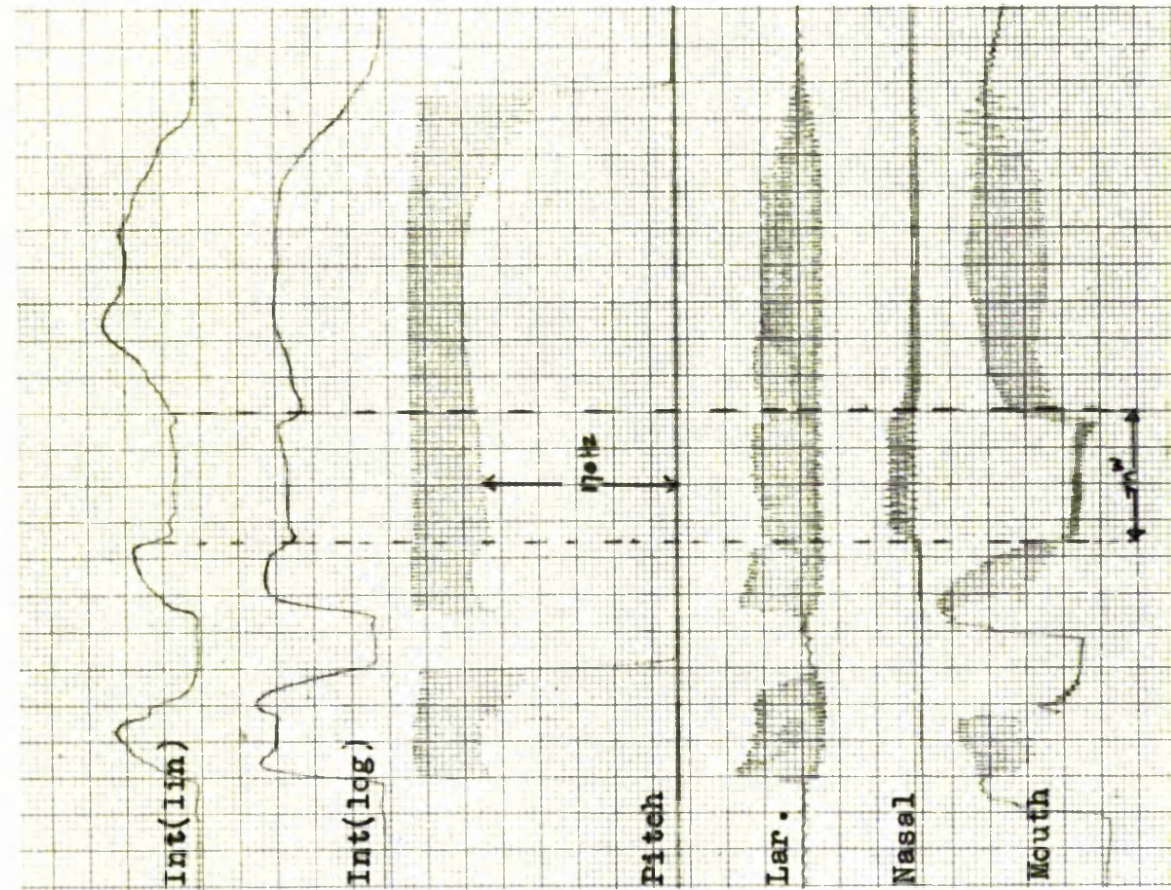
r

i

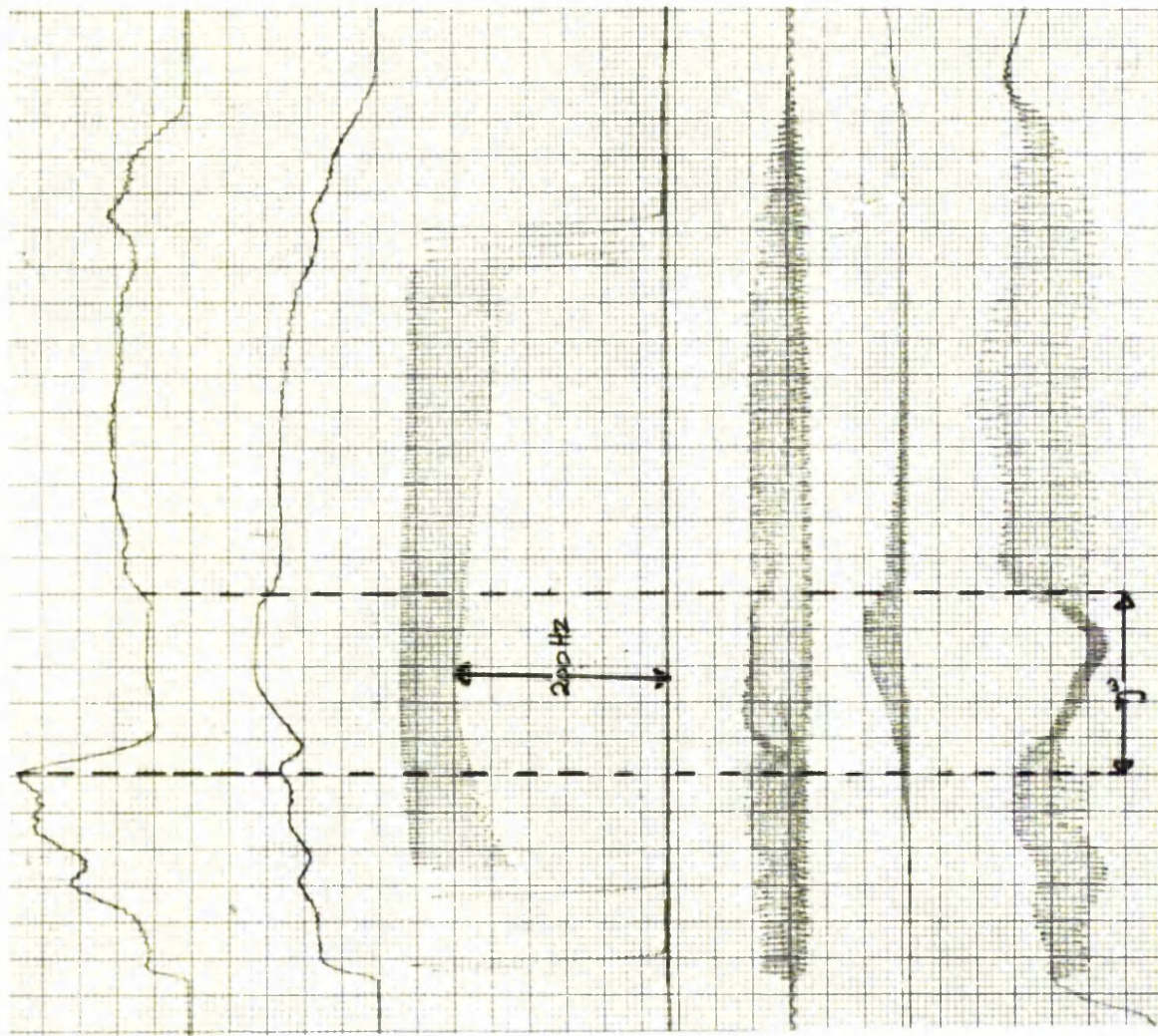
y

b.

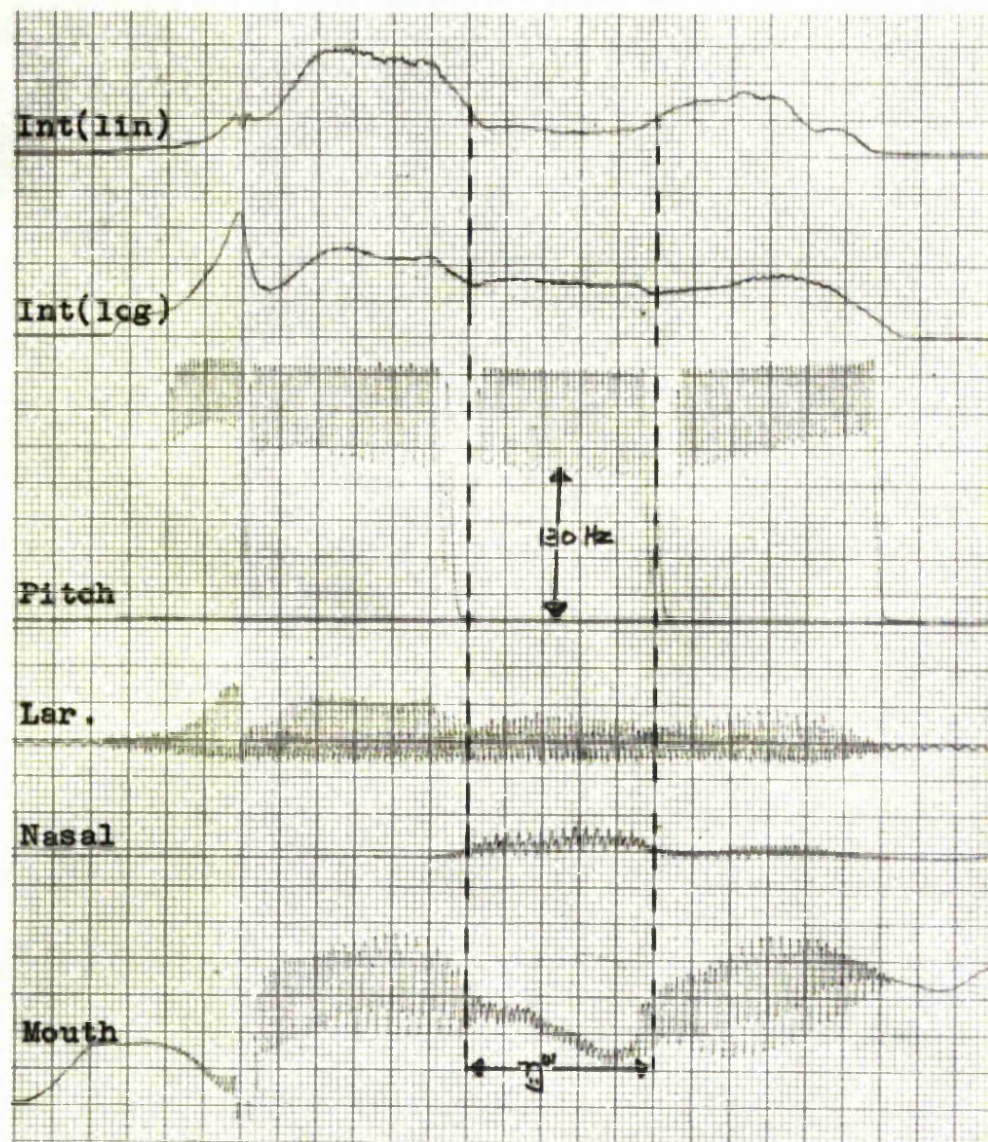
m.



Mingogram 20. [kùtì m'í] "ideo. of silence"



Mingogram 21. [wáj'íwà] "it (beer) has been drunk"

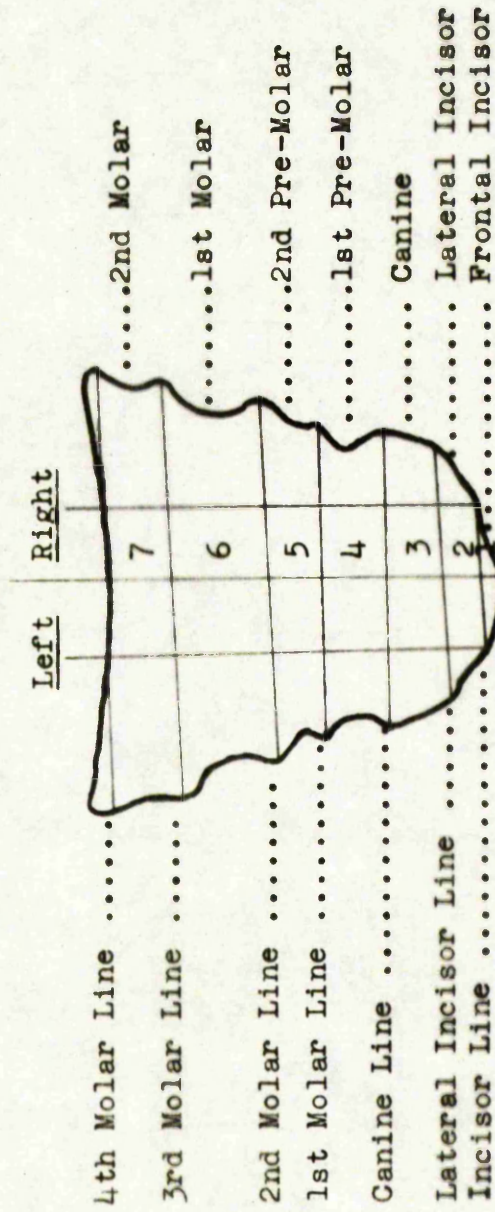


Mingogram 22. [dòŋ'ɛ] "drop of rain" or "water"

The Palatogram Figure (b)

Mr. A. Pongweni

Zones



Mr. A. Pongweni

